

*DEV*

*John Thomas*

NWC TP 5594

*C.I*

# Difference Equations and the Optimal Control of Single Server Queueing Systems

by

Frank C. Reed

*Research Department*

JUNE 1974

Approved for public release; distribution unlimited.

19991214 082

## Naval Weapons Center

CHINA LAKE, CALIFORNIA 93555



# Naval Weapons Center

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

Paul E. Pugh, RAdm., USN ..... Commander  
Leroy Riggs ..... Technical Director (Acting)

## FOREWORD

This report was written by Mr. Reed in the course of pursuing a Ph. D. during calendar years 1972-1974 at Stanford University. The work was directed by Professor Frederick S. Hillier, Department of Operations Research, and it has been published as Department of Operations Research Report No. 44, March 22, 1974.

This report gives examples of the use of difference equations in the investigation of optimal control policies in queueing systems. It is the second in a series of four reports which will cover Mr. Reed's research at Stanford in queueing optimization. Since queueing systems and their optimization have applications to a number of DOD problems the report is reproduced here as an NWC TP for distribution to various defense installations.

Mr. Reed was supported by Navy Director of Laboratory Programs, Task Assignment R00001-R01405. Review and report preparation at Stanford University were supported in part by National Science Foundation Grant GK-35491 and Army and Navy Contract N00014-67-A-0112-0052 (NR-042-002).

Released by  
DELBERT E. ZILMER, Head  
*Mathematics Division*  
3 June 1974

Under authority of  
HUGH W. HUNTER, Head  
*Research Department*

NWC Technical Publication 5594

Published by . . . . . Research Department  
Collation . . . . . Cover, 52 leaves  
First printing. . . . . 150 unnumbered copies

## UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NWC TP 5594	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) DIFFERENCE EQUATIONS AND THE OPTIMAL CONTROL OF SINGLE SERVER QUEUEING SYSTEMS		5. TYPE OF REPORT & PERIOD COVERED A research report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s)  Frank C. Reed		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS  Naval Weapons Center China Lake, CA 93555		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  1152N, R00001, R01405 160070-6
11. CONTROLLING OFFICE NAME AND ADDRESS  Naval Weapons Center China Lake, CA 93555		12. REPORT DATE June 1974
		13. NUMBER OF PAGES 102
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)  UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Single-Server Queueing Systems, Optimal Control, Expected Discounted Costs, Expected Average Cost per Unit Time, Intermittent Service, Service Rates Selection, Bulk Service, Difference Equations		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  See reverse side		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

(U) *Difference Equations and the Optimal Control of Single Server Queueing Systems* by Frank C. Reed. China Lake, Calif., Naval Weapons Center, June 1974. 102 pp. (NWC TP 5594, publication UNCLASSIFIED).

(U) This report demonstrates the use of difference equations in solving optimal control problems in single server queueing systems. One obtains the discounted or relative cost function associated with a specific stationary policy by solving an appropriate system of difference equations. The policy improvement algorithm is applied parametrically leading to a characterization of the cost function satisfying the functional equation of optimality. If this cost function satisfies an appropriate sufficient condition, the associated stationary policy is optimal.

(U) The method of solution is illustrated by solving three queueing optimization problems. These problems include optimal control of the M/G/1 queue with intermittent service, a bulk queueing version of this same problem, and control of the M/M/1 queue with selection of running speed. All of these problems have been investigated by other authors. Results in this report believed to be new include a complete characterization of optimal policies for the optimal control of the M/G/1 queue in the discounted case, the extension of the optimal control of the bulk queueing problem from instantaneous to general service, and the determination of an optimal speed selection policy for the M/M/1 queue without solving a sequence of truncated problems.

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

CONTENTS

Introduction. . . . .	1
The M/G/1 Queue with Removable Server . . . . .	6
Existence of Stationary Optimal Policy for the Average Cost Case . . . . .	7
Qualitative Attributes of an Optimal Policy for the Average Cost Case . . . . .	11
Quantitative Results Associated with an Optimal Policy in the Average Cost Case . . . . .	17
Qualitative and Quantitative Results for the Discounted Case .	22
Optimal Control of a Bulk Queueing System . . . . .	46
Existence of a Stationary Optimal Policy . . . . .	48
Qualitative Attributes of an Optimal Policy. . . . .	53
Determination of an Optimal Policy . . . . .	55
The M/M/1 Queue with Variable Service Rate. . . . .	66
Existence of a Stationary Optimal Policy . . . . .	66
Qualitative Attributes of an Optimal Policy. . . . .	70
Quantitative Results for the Linear Holding Cost Case. . . . .	76
Appendixes:	
A. Functional Equations of Optimality. . . . .	83
B. Glossary. . . . .	93
References. . . . .	100

NWC TP 5594

DIFFERENCE EQUATIONS AND THE OPTIMAL  
CONTROL OF SINGLE SERVER  
QUEUEING SYSTEMS

by

Frank C. Reed  
Stanford University and Naval Weapons Center, China Lake

1. Introduction

This report describes, by way of example, the use of difference equations to obtain optimal control policies in single server queueing systems. The difference equations solved in this report give explicit expressions for expected discounted costs or relative costs as defined in Howard [5] for use in a policy improvement algorithm to solve the dynamic programming functional equations of optimality. Assuming that a stationary optimal policy exists the solution is carried out in the following way:

- (i) Determine the qualitative attributes of an optimal stationary policy, thus restricting the family of stationary policies and class of difference equations one must consider.
- (ii) If necessary, apply the policy improvement algorithm and determine the difference equation solution to the functional equations of optimality.
- (iii) Show that the policy obtained under (ii) satisfies sufficient conditions for optimality.

A classical method of solving queueing optimization problems is to concentrate on (i) without consideration of the difference equations involved. Once the family of stationary policies has been suitably restricted, one may then construct specialized algorithms for arriving at an optimal policy. Since it is possible to infer qualitative properties of optimal policies by consideration of (ii), the difference equation approach allows for flexibility in the combined use of (i) and (ii) in determining optimal queueing policies. It appears, moreover, that the solution of the difference equations involved may lead to highly efficient computing algorithms.

To illustrate the method, optimal control policies are obtained for three single server queueing systems. Aside from the difference equation approach the solutions depend on material presented in Reed [7]. That report establishes sufficient conditions for both the existence of stationary optimal policies and the optimality of stationary policies in Markov decision processes for which an assumption of bounded costs is not appropriate.

In the discounted cost case it is shown that if costs are non-negative, a stationary optimal policy exists. Sufficient conditions for the optimality of a stationary policy require that an explicit expression for the difference equations associated with the policy be available. Briefly in the average cost case it is shown that if

A(1): There is a policy for which the average cost per unit time is finite,

A(2): There exists a state that is positive recurrent over all policies,

A(3): Relative costs for all policies are bounded below, then a stationary optimal policy exists. Sufficient conditions for optimality

of a stationary policy require explicit expressions for relative costs obtained by solving the appropriate difference equations.

It will be assumed that the reader is familiar with the assumptions, definitions, notation, and precise form of the above results as presented in Reed [7]. For reference, a summary of basic results and notation in that report is given in Appendix B.

Section 2 considers the control of an M/G/1 queue with removable server, when the optimization criteria are the minimum expected average cost per unit time and the minimum expected discounted cost over an infinite horizon. The average cost case has been studied by Heyman [4] and Sobel [10]. The solution presented here for the average cost case provides a new proof for the existence of a stationary optimal policy. Proofs different from Heyman's are given to restrict the class of stationary policies in which an optimal policy lies, and difference equations are solved to express the average cost per unit time as a function of two parameters. This function is easily minimized and Heyman's results are extended to include rewards and a more general holding cost assumption. For the discounted case investigated by Heyman [4], Bell [1], and Blackburn [2], the stationary optimal policy is obtained using the difference equation approach to solve explicitly the functional equation associated with optimality. It is then shown that this solution satisfies sufficient conditions for optimality. The final result is a complete characterization of all optimal policies, without resorting to numerical application of the policy improvement algorithm or stopping rule algorithms.

Section 3 presents two versions of a bulk queueing problem which are both generalizations of the mail truck problem presented by Ross [8, pp. 164 ff]. This problem is solved for the average cost case and allows a general

distribution for the time to perform bulk service. The two versions result from different cost structure assumptions. For both problems it is shown that a stationary optimal policy exists, and difference equations are used to construct a policy improvement algorithm. In addition, calculations made in Section 4 consider the control of the M/M/1 queue with variable service rate that was originally investigated by Crabbill [3], Kakalik [6], Sabeti [9], and more recently for closed systems by Torbett [11]. For Crabbill's form of this problem it is shown that a stationary optimal policy exists and that it has a simple form. Unlike Crabbill's proof, this proof avoids truncation of the original problem. Using difference equations the expected average cost per unit time is obtained as a function of a parameter which describes the family of permissible stationary policies. A method of determining the minimum of this function is presented.

## 2. The M/G/1 Queue with Removable Server

Consider the situation where an M/G/1 queue is controlled by turning the server off and on. Customers arrive according to a Poisson process with rate  $\lambda > 0$ . Service times are non-negative, independent random variables with common distribution function  $B$ . It is assumed that the mean service time  $\mu^{-1} = \int t dB(t)$  is finite and positive. The probability of finding the server off at time  $t$  is denoted by  $\alpha(t)$ , and the probability of finding the server on at time  $t$  is denoted by  $\beta(t)$ .

With the arrival of a customer, the server begins service and the system enters a state of service. After a time  $t$ , the server is turned off and the system enters a state of idleness. The time between the start of service and the end of service is denoted by  $S(t)$ . The probability of finding the system in state  $i$  at time  $t$  is denoted by  $P_i(t)$ .

and

$$\rho = \lambda\mu^{-1} < 1.$$

There is a cost  $R_1$  of turning the server on when he is idle, and a cost  $R_2$  of turning the server off when he is active. There also is a cost  $r_1$  per unit time of maintaining the queueing system when it is idle, and a cost  $r_2$  per unit time of operating the system when the server is active. In addition, there is a holding cost of  $h$  per customer per unit time. For the average cost case there is a reward  $G$  received at the completion of a service.<sup>1</sup> Rewards are not included in the discounted case.

Decisions are made at the time of service completion, or if the server is idle decisions are made at the time of customer arrival. In the former case, the server may either remain active and continue service or he may shut down and go idle. In the latter case he may either remain idle or start up. We associate  $k = 0$  with the decision to remain idle or shut down. We associate  $k = 1$  with the decision to remain active or start up.

The state of the system is described by the pair  $(i, j)$  where  $i$  indicates the number of customers in the queue (waiting or in service);  $j = 0$  if the server is idle and  $j = 1$  if the server is active. We shall adopt the convention that for any realization of the stochastic process associated with this queueing system  $i$  is right continuous and  $j$  is left continuous with respect to the time parameter.

## 2.1. Existence of Stationary Optimal Policy for the Average Cost Case

We now proceed to verify assumptions A(1), A(2), and A(3) of Reed [7] for the existence of a stationary optimal policy. At this point we must

---

<sup>1</sup> Rewards were not considered in Heyman's work and this provides a minor extension to his results.

place some restrictions on permissible service distributions.

LEMMA 2.1.

If  $\mu_2 = \int t^2 dB(t) < \infty$ , then A(1) is satisfied.

Proof. We consider that stationary policy  $f_0$  which always provides service. It follows from Reed [7] and Appendix A that the difference equations associated with this policy are

$$\begin{aligned} v_i - \sum_{k=0}^{\infty} v_{i+k-1} \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \phi_{f_0} \mu^{-1} \\ = (h_i + r_2) \mu^{-1} + \frac{\lambda h \mu_2}{2} - G \quad i = 1, 2, \dots \\ v_0 - v_1 + \phi_{f_0} / \lambda = r_2 / \lambda \end{aligned} \tag{2.1}$$

where  $v_0 = 0$  allows determination of  $\phi_{f_0}$ .

Setting

$$v_i = b_i + c i^2,$$

substituting in (2.1), and interchanging summation and integration which is permissible in this case, we have

$$c = \frac{h \mu^{-1}}{2(1-\rho)}$$

$$b = \frac{h \mu^{-1}}{2(1-\rho)} + \frac{\lambda h \mu_2}{2(1-\rho)^2} - \frac{[(\phi_{f_0} - r_2) \mu^{-1} + G]}{1-\rho}.$$

Since

$$\begin{aligned}\phi_{f_0} &= \lambda v_1 + r_2 = \lambda(b+c) + r_2, \\ \phi_{f_0} &= r_2 + hp + \frac{\lambda^2 h \mu_2}{2(1-p)} - \lambda G.\end{aligned}\tag{2.2}$$

Since  $\mu_2 < \infty$  we may set  $M$  in A(1) equal to  $\phi_{f_0}$ .

With respect to the optimal cost function  $V$  we have

$$V((i,0)) \leq v_{f_0}((i,0)) = R_1 + bi + ci^2,$$

$$V((i,1)) \leq v_{f_0}((i,1)) = bi + ci^2,$$

$$\phi \leq \phi_{f_0}.$$

LEMMA 2.2. Assumption A(2) is satisfied.

Proof. We must show there exists a state recurrent over all policies  $\pi$  for which  $\phi_\pi \leq \phi_{f_0}$ . We define

$$I_0 = \{(i,0): h_i + r_1 \leq \phi_{f_0} + \lambda G\} \equiv \{(0,0) \dots (i_0,0)\}$$

$$I_1 = \{(i,1): h_i + r_2 \leq \phi_{f_0} + \lambda G\} \equiv \{(0,1) \dots (i_1,1)\}.$$

We note that  $\phi_\pi + \lambda G$  is the cost per unit time of any policy  $\pi$  which services customers in such a way that the number of customers held without service does not grow without bound. For any such policy  $\pi$  for which

$$\phi_{\pi} + \lambda G \leq \phi_{f_0} + \lambda G$$

the set  $I_0 \cup I_1$  is recurrent. Since  $1/\lambda < \infty$ , any policy  $\pi$  will consist of a sequence of services possibly separated by idle periods of finite duration. Moreover, since  $\mu^{-1} < \infty$ ,  $\pi$  will give rise to a sequence of services with finite expected times between service. With each service completion the probability of a transition into  $\overline{I_0 \cup I_1}$  (the complement of  $I_0 \cup I_1$ ) exceeds

$$\sum_{k=k_0}^{\infty} \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) > 0$$

where  $k_0 = \max [i_0 + 1, i_1 + 1]$ . It follows that  $\overline{I_0 \cup I_1}$  is recurrent. In going from  $\overline{I_0 \cup I_1}$  to  $I_0 \cup I_1$  the state  $(\max [i_0 + 1, i_1 + 1], 1)$  must be entered; hence this state is recurrent over all  $\pi$  for which  $\phi_{\pi} \leq \phi_{f_0}$ .

LEMMA 2.3. If  $\mu_2 < \infty$ , A(3) holds.

Proof. To prove this lemma we use Reed [7], Theorem 4.3 and Appendix A.

We have

$$c_{(i,0)}^{(0)} = \frac{hi + r_1}{\lambda}$$

$$c_{(i,0)}^{(1)} = R_1$$

$$c_{(i,1)}^{(1)} = (hi + r_2)\mu^{-1} + \frac{\lambda h \mu_2}{2} - G$$

$$c_{(i,1)}^{(0)} = R_2$$

$$t_{(i,0)}^{(0)} = \frac{1}{\lambda}$$

$$t_{(i,0)}^{(1)} = 0$$

$$t_{(i,1)}^{(1)} = \mu^{-1}$$

$$t_{(i,1)}^{(0)} = 0.$$

Since  $t_{(i,0)}^{(1)} = t_{(i,1)}^{(0)} = 0$ , we must verify that the elimination of trivial sequences as defined in Reed [7], Section 3.2.2 implies that Condition 1 holds. All trivial sequences have non-negative costs associated with them and the average cost per unit time cannot be reduced by including them. Consequently, we may assume that every instantaneous action is followed by an action with positive expected transition time. It follows that Condition 1 is satisfied.

Since  $\mu_2 < \infty$ ,  $\phi_{f_0} < \infty$  and the set  $S$  of Theorem 4.3 is finite, it follows that  $V$  is bounded below since  $C_i(k)$  are bounded below.

We summarize these results in a theorem.

#### THEOREM 2.4.

If  $\mu_2 < \infty$ , there exists a stationary optimal policy.

#### 2.2. Qualitative Attributes of an Optimal Policy for the Average Cost Case

If  $f^*$  is an optimal stationary policy, then the cost function  $V_{f^*}$  has certain properties which we shall now investigate. We recall the functional equations which  $V_{f^*}$  must satisfy

$$V_{f^*}(i,0) = \min [V_{f^*}(i+1,0) - \frac{\phi_{f^*}}{\lambda} + \frac{h_i + r_1}{\lambda}, V_{f^*}(i,1) + R_1] \quad i = 0, 1, 2, \dots$$

$$v_{f^*}(i,1) = \min \left[ \sum_{k=0}^{\infty} v_{f^*}(i+k-1,1) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) \right]$$

$$-\phi_{f^*}^{-1} + (hi + r_2)\mu^{-1} + \frac{\lambda h \mu_2}{2} - G,$$

$$v_{f^*}(i,0) + R_2] \quad i = 1, 2, \dots$$

$$v_{f^*}(0,1) = \min [v_{f^*}(1,1) - \frac{\phi_{f^*}}{\lambda} + \frac{r_2}{\lambda}, v_{f^*}(0,0) + R_2].$$

With respect to these functional equations we prove the following lemma.

LEMMA 2.5. If  $R_1 + R_2 > 0$ , then for each  $i$  there are only three possibilities,

$$(i) \quad v_{f^*}(i,1) = v_{f^*}(i,0) + R_2 \quad i = 0, 1, 2, \dots$$

$$v_{f^*}(i,0) = v_{f^*}(i+1,0) - \frac{\phi_{f^*}}{\lambda} + \frac{hi + r_1}{\lambda} \quad i = 0, 1, \dots$$

$$(ii) \quad v_{f^*}(i,1) = \sum_{k=0}^{\infty} v_{f^*}(i+k-1,1) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t)$$

$$-\phi_{f^*}^{-1} + (hi + r_2)\mu^{-1} + \frac{\lambda h \mu_2}{2} - G \quad i = 1, 2, \dots$$

$$v_{f^*}(0,1) = v_{f^*}(1,1) - \frac{\phi_{f^*}}{\lambda} + \frac{r_2}{\lambda}$$

$$v_{f^*}(i,0) = v_{f^*}(i+1,0) - \frac{\phi_{f^*}}{\lambda} + \frac{hi + r_1}{\lambda} \quad i = 0, 1, 2, \dots$$

$$(iii) \quad v_{f^*}(i,0) = v_{f^*}(i,1) + R_1 \quad i = 0, 1, \dots$$

$$v_{f^*}(i,1) = \sum_{k=0}^{\infty} v_{f^*}(i+k-1,1) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t)$$

$$- \phi_{f^*}^{-1} + (hi + r_2)\mu^{-1} + \frac{\lambda h \mu_2}{2} - G \quad i = 1, 2, \dots$$

$$v_{f^*}(0,1) = v_{f^*}(1,1) - \frac{\phi_{f^*}}{\lambda} + \frac{r_2}{\lambda}.$$

Proof. The only other possibility is

$$v_{f^*}(i,1) = v_{f^*}(i,0) + R_2$$

$$v_{f^*}(i,0) = v_{f^*}(i,1) + R_1$$

so that

$$0 = R_1 + R_2,$$

a contradiction.

We note that if (i) in Lemmas 2.5 holds, then

$$f^*(i,0) = 0$$

and

$$f^*(i,1) = 0.$$

Any  $i$  for which (i) in Lemma 2.5 holds is called an idle integer. If (ii) in Lemma 2.5 holds, then

$$f^*(i,0) = 0$$

$$f^*(i,1) = 1.$$

Any  $i$  for which this holds is called an indifference integer. Finally, if (iii) in Lemma 2.5 holds, then

$$f^*(i,0) = 1$$

$$f^*(i,1) = 1$$

and  $i$  will be called a service integer. As a consequence, if  $i$  is the number of customers in the queue, then the determination of an optimal policy  $f^*$  is equivalent to categorizing all integers  $i$  as idle, indifference, or service.

LEMMA 2.6. The set of  $i$  for which (i) in Lemma 2.5 holds is bounded for  $f^*$ .

Proof. Assume the set of  $i$  for which (i) holds is unbounded and let  $N$  be such that (i) holds and  $hN + \min(r_1, r_2) > \phi_{f_0} + \lambda G$ . Also let  $S_0 = \{(i,1): i \geq N\}$ . If service is never performed,  $\phi \rightarrow \infty > \phi_{f_0}$ , a contradiction.

Thus, service periods alternate with finite idle periods and  $S_0$  is accessible in finite time. Once  $S_0$  is entered the number in the queue never drops below  $N$ .

It follows that

$$\phi + \lambda G > hN + \min(r_1, r_2) > \phi_{f_0} + \lambda G,$$

a contradiction. Hence the set of  $i$  for which (i) holds is bounded.

LEMMA 2.7. If  $R_1 + R_2 > 0$ ,

$$-R_1 \leq v_f^*(i,1) - v_f^*(i,0) \leq R_2$$

with one equality holding when case (i) or (iii) of Lemma 2.5 holds.

Proof. If (i) in Lemma 2.5 holds,

$$v_f^*(i,1) - v_f^*(i,0) = R_2.$$

If (iii) in Lemma 2.5 holds,

$$v_f^*(i,1) - v_f^*(i,0) = -R_1.$$

If (ii) holds, then from the functional equations defining  $v_f^*$ ,

$$v_f^*(i,0) \leq v_f^*(i,1) + R_1,$$

$$v_f^*(i,1) \leq v_f^*(i,0) + R_2,$$

from which the desired result follows.

LEMMA 2.8. There exists an  $N$  such that (iii) of Lemma 2.5 holds for all  $i \geq N$ .

Proof. From Lemma 2.6 it follows that for  $i$  sufficiently large

$$\begin{aligned} v_{f^*}(i,1) &= \sum_{k=0} v_{f^*}(i+k-1,1) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) \\ &\quad - \phi_{f^*}^{-1} + (h_i + r_2) \mu^{-1} + \frac{\lambda h \mu_2}{2} - G. \end{aligned}$$

It follows from Lemma 2.1 that

$$v_{f^*}(i+1,1) - v_{f^*}(i,1) = \frac{h \mu^{-1}}{1-\rho} i + \frac{h \mu^{-1}}{1-\rho} + \frac{\lambda h \mu_2}{2(1-\rho)^2} + \frac{(r_2 - \phi_{f^*}) \mu^{-1} - G}{(1-\rho)}.$$

Now assume (ii) in Lemma 2.5 holds.

$$v_{f^*}(i+1,0) - v_{f^*}(i,0) = \frac{\phi_{f^*} - r_1}{\lambda} - \frac{h}{\lambda} i$$

so that

$$\begin{aligned} v_{f^*}(i+1,1) - v_{f^*}(i+1,0) + [v_{f^*}(i,0) - v_{f^*}(i,1)] \\ = \frac{h \mu^{-1} i}{\rho(1-\rho)} + \frac{h \mu^{-1}}{1-\rho} + \frac{\lambda h \mu_2}{2(1-\rho)^2} - \frac{\phi_{f^*} \mu^{-1}}{\rho(1-\rho)} + \frac{r_2 \mu^{-1}}{1-\rho} + \frac{r_1 \mu^{-1}}{\rho} - \frac{G}{1-\rho}. \end{aligned}$$

From Lemma 2.7 we have

$$R_1 + R_2 \geq \frac{h \mu^{-1} i}{\rho(1-\rho)} + k.$$

For  $i$  sufficiently large this leads to a contradiction.

Now applying Lemma 2.6 we let  $N_1$  be the maximum integer for which (i) in Lemma 2.5 holds. If there is no idle integer we set  $N_1 = -1$ .

Applying Lemma 2.8 we let  $N_2$  be the smallest integer greater than or equal to  $N_1$  for which (iii) in Lemma 2.5 holds. From Lemma 2.5 it follows that  $N_1 < N_2$  and  $i$  with  $N_1 < i < N_2$  are indifference integers. We shall always be concerned with a well-defined set of closed states which communicate with  $(N_1, 1)$  or  $(0, 1)$  if  $N_1 = -1$ .

Since  $(N_2 + j, 0)$  for  $j > 0$  are transient with respect to this closed set, we may specify that decision 1 is made in states  $(N_2 + j, 0)$ . Thus all  $i \geq N_2$  may be regarded as service integers. Since the states  $(N_1 - j, 0)$  and  $(N_1 - j, 1)$  for  $j > 0$  are transient with respect to this closed set, we may arbitrarily specify that decision 0 is made in these states. Thus all  $i \leq N_1$  may be regarded as idle integers. We summarize these remarks in the following theorem:

**THEOREM 2.9.**

A stationary optimal policy is characterized by two integers,  $N_1$  and  $N_2$  with  $-1 \leq N_1 < N_2$  where all  $i \leq N_1$  are idle integers, all  $i \geq N_2$  are service integers, and all  $i$  with  $N_1 < i < N_2$  are indifference integers.

**2.3. Quantitative Results Associated with an Optimal Policy in the Average**

**Cost Case**

We proceed to determine optimal values for  $N_1$  and  $N_2$ .

**LEMMA 2.10.** Let  $f(N_1, N_2)$  be a stationary policy with  $N_1$  and  $N_2$  defined in Theorem 2.9. Then the average cost per unit time associated with  $f$  is given by

$$\begin{aligned}\phi_{f(N_1, N_2)} &= r_1(1 - \rho) + r_2\rho + h(\rho + \frac{\lambda^2 \mu_2}{2(1 - \rho)}) \\ &+ \frac{(N_2 - 1 + N_1)h}{2} + \frac{\lambda(1 - \rho)(R_1 + R_2)}{N_2 - N_1} - \lambda G \quad (2.3)\end{aligned}$$

for  $N_1 \geq 0$ , and

$$\phi_{f(-1, N_2)} = \phi_{f_0} .$$

Proof. If  $N_1 = -1$ , the server will start in finite time and then never be turned off so  $\phi_{f(-1, N_2)} = \phi_{f_0}$ . Dropping reference to  $N_1$  and  $N_2$  we have the following system to solve for  $N_1 \geq 0$ ,

$$V(i, 0) = V(i + 1, 0) - \frac{\phi}{\lambda} + \frac{hi + r_1}{\lambda} \quad 0 \leq i \leq N_2 - 1$$

$$V(i, 1) = \sum V(i + k-1, 1) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t)$$

$$+ (hi + r_2)^{\mu-1} + \frac{\lambda h \mu_2}{2} - \phi^{\mu-1} - G \quad i = N_1 + 1 \dots$$

$$V(i, 1) = V(i, 0) + R_2 \quad i = 0, 1, \dots, N_1$$

$$V(i, 0) = V(i, 1) + R_1 \quad i = N_2, N_2 + 1, \dots$$

We set  $V(0, 0) = 0$  and let

$$V(i, 0) = B_0 i + C_0 i^2,$$

with the result

$$B_0 = \frac{\phi - r_1}{\lambda} + \frac{h}{2\lambda}$$

$$C_0 = - \frac{h}{2\lambda} .$$

Setting

$$V(i,1) = A_1 + B_1 i + C_1 i^2$$

$$B_1 = \frac{h\mu^{-1}}{2(1-\rho)} + \frac{(r_2 - \phi)\mu^{-1} - G}{1-\rho} + \frac{\lambda h\mu_2}{2(1-\rho)^2}$$

$$C_1 = \frac{h\mu^{-1}}{2(1-\rho)} .$$

Since

$$V(N_2,0) = V(N_2,1) + R_1$$

and

$$V(N_1,0) = V(N_1,1) - R_2,$$

we have

$$N_1(B_0 - B_1) - A_2 = (C_1 - C_0)N_1^2 - R_2$$

$$N_2(B_0 - B_1) - A_2 = (C_1 - C_0)N_2^2 + R_1$$

and

$$B_0 - B_1 = \frac{h}{2\lambda(1-\rho)} \cdot \frac{(N_2^2 - N_1^2)}{N_2 - N_1} + \frac{R_1 + R_2}{N_2 - N_1}.$$

Since

$$B_0 - B_1 = \frac{\phi}{\lambda(1-\rho)} - \frac{r_1}{\lambda} - \frac{r_2 \mu^{-1}}{1-\rho} - \frac{\lambda h \mu_2}{2(1-\rho)^2} + \frac{h(1-2\rho)}{2\lambda(1-\rho)} + \frac{G}{1-\rho},$$

(2.3) follows.

#### THEOREM 2.11.

A stationary optimal policy is characterized by a single integer  $N$ . If  $N = 0$ , all  $i$  are service integers and  $f_0$  is optimal. If  $N \geq 1$ ,  $i = 0$  is an idle integer, all  $i \geq N$  are service integers, and  $i$  for which  $0 < i < N$  are indifference integers.

Proof. For stationary policies described by  $N_1 \geq 0$  and  $N_2$ , we have

$$\frac{d\phi_{f(N_1, N_2)}}{dN_1} = \frac{h}{2} + \frac{\lambda(1-\rho)(R_1 + R_2)}{(N_2 - N_1)^2} > 0$$

and  $\phi_{f(N_1, N_2)}$  is minimized by setting  $N_1 = 0$ . It is easily shown that

$\phi_{f(0, N_2)}$  is convex in  $N_2$  and the optimal value of  $N_2$  is one of the two integers adjacent to

$$\hat{N}_2 = \left[ \frac{2\lambda(1-\rho)(R_1 + R_2)}{h} \right]^{1/2}$$

Defining  $\langle a \rangle$  to be the smallest integer greater than or equal to  $a$ , the minimum value of  $\phi_{f(N_1, N_2)}$  is given by

$$\phi_{f(N_1, N_2)} = \min [\phi_{f_0}, \phi_{f(0, \langle \hat{N}_2 \rangle - 1)}, \phi_{f(0, \langle \hat{N}_2 \rangle)}].$$

If the minimum is attained for the first entry,  $N = 0$ . If the minimum is attained by the second or third entry, then  $N = \langle \hat{N}_2 \rangle - 1$  or  $N = \langle \hat{N}_2 \rangle$ , respectively.

These results are easily extended to include different holding rates,  $h_1$  when the server is idle and  $h_2$  when active. In this case

$$\begin{aligned} \phi_{f(N_1, N_2)} &= r_1(1-\rho) + r_2\rho + h_2\rho + \frac{\lambda^2 h_2 \mu_2}{2(1-\rho)} \\ &+ \frac{(N_2 - 1 + N_1)}{2} (h_2\rho + h_1(1-\rho)) + \frac{\lambda(1-\rho)(R_1 + R_2)}{N_2 - N_1} \end{aligned}$$

$$\hat{N}_1 = 0,$$

and

$$\hat{N}_2 = \left[ \frac{2\lambda(1-\rho)(R_1 + R_2)}{h_2\rho + h_1(1-\rho)} \right]^{1/2}.$$

2.4. Qualitative and Quantitative Results for the Discounted Case

Since  $G = 0$ ,  $C_i(k) \geq 0$ , and it follows from Reed [7] Theorem 3.12 that there exists a stationary  $\beta$ -optimal policy. Our plan is to determine a  $\beta$ -optimal improvement policy as defined in Reed [7] Section 3.2.2 and then impose sufficient conditions that this policy be optimal. The final result is a complete characterization of all optimal policies, without resorting to numerical application of policy improvement or stopping rule algorithms. As a consequence both qualitative and quantitative results will be obtained together.

We define  $v_\beta(i,j)$  as the total expected discounted cost over an infinite horizon given the process begins in state  $(i,j)$ . From Reed [7] and Appendix A, it follows that the functional equations associated with an optimal stationary policy are as follows:

$$v_\beta(i,0) = \min\left\{\frac{\lambda}{\lambda+\beta}v_\beta(i+1,0) + (h_i + r_1)/(\lambda+\beta), v_\beta(i,1) + R_1\right\}, \quad i \geq 0$$

$$v_\beta(0,1) = \min\{v_\beta(0,0) + R_2, \frac{\lambda}{\lambda+\beta}v_\beta(1,1) + \frac{r_2}{\lambda+\beta}\}$$

$$v_\beta(i,1) = \min\{v_\beta(i,0) + R_2,$$

$$\sum_{k=0}^{\infty} v_\beta(i+k-1,1) \int \frac{(\lambda t)^k}{k!} e^{-(\lambda+\beta)t} dB(t) + (h_i + r_2) \frac{(1-\tilde{B}(\beta))}{\beta} \\ + \frac{\lambda h}{\beta^2} (1-\tilde{B}(\beta) - \beta \int t e^{-\beta t} dB(t))\}, \quad i \geq 1,$$

$$\text{where } \tilde{B}(\beta) = \int_0^\infty e^{-\beta t} dB(t).$$

We shall refer to the following equations in characterizing integers associated with various types of policies:

$$v_{\beta}(i,0) - \frac{\lambda}{\lambda+\beta} v_{\beta}(i+1,0) = \frac{hi + r_1}{\lambda+\beta} \quad i \geq 0 \quad (2.4)$$

$$v_{\beta}(0,1) - \frac{\lambda}{\lambda+\beta} v_{\beta}(1,1) = \frac{r_2}{\lambda+\beta} \quad (2.5)$$

$$\begin{aligned} v_{\beta}(i,1) &= \sum_k v_{\beta}(i+k-1,1) \int \frac{(\lambda t)^k}{k!} e^{-(\lambda+\beta)t} dB(t) \\ &= (hi + r_2) \frac{(1-B(\beta))}{\beta} + \frac{\lambda h}{\beta^2} (1-B(\beta) - \beta \int te^{-\beta t} dB(t)) \quad i \geq 1 \end{aligned}$$

$$v_{\beta}(i,0) = v_{\beta}(i,1) + R_1 \quad (2.6)$$

$$v_{\beta}(i,1) = v_{\beta}(i,0) + R_2 . \quad (2.7)$$

Any  $i$  for which (2.4) and (2.7) hold is defined as an idle integer. Any  $i$  for which (2.5) and (2.6) hold is defined as a service integer, and any  $i$  for which (2.4) and (2.5) hold is defined as an indifference integer.

Since  $C_i(k) \geq 0$ , the elimination of trivial sequences implies that there are no integers for which (2.6) and (2.7) hold.

We consider solutions to (2.4) and (2.5) of the form

$$v_{\beta}(i,0) = A_1 + B_1 i + K_1 \tilde{A}(\beta)^{-i}$$

and

$$v_{\beta}(i,1) = A_2 + B_2 i + K_2 \tilde{G}(\beta)^{-i},$$

where

$$\tilde{A}(\beta) = \frac{\lambda}{\lambda + \beta}$$

and  $\tilde{G}(\beta)$  satisfies

$$\tilde{G}(\beta) = \tilde{B}(\beta + \lambda(1 - \tilde{G}(\beta))).$$

Defining

$$\tilde{H}(\beta) = \frac{h\tilde{B}(\beta)}{\beta(1 - \tilde{B}(\beta))}$$

and

$$\psi_1 = \frac{r_2 - r_1}{\beta} + R_1,$$

the solution for (2.4) and (2.5) (except for  $i = 0$ ) is specified by

$$A_1 = r_1/\beta + \lambda h/\beta^2$$

$$A_2 = r_2/\beta + \lambda h/\beta^2 - \tilde{H}(\beta)$$

$$B_1 = B_2 = h/\beta,$$

and  $K_1$  and  $K_2$  are determined by boundary conditions that hold. Since  $\beta$  is fixed, we shall set  $\tilde{A}(\beta) = A$ ,  $\tilde{G}(\beta) = G$ , and  $\tilde{H}(\beta) = H$ .

THEOREM 2.12.

If  $H \leq \psi_1 \leq HA(1-G)/(1-AG) + R_1 + R_2$ , that policy for which all  $i \geq 0$  are indifference integers is a  $\beta$ -optimal improvement policy.

Proof. In this case we solve (2.4) and (2.5) for all  $i$ , where the equation in  $V_\beta(0,1)$  imposes a boundary condition from which  $K_2$  is determined. Imposing this boundary condition gives

$$K_2 = H(1-A)/(1-AG)$$

and, since there is no boundary condition on  $K_1$ , we set  $K_1 = 0$ . The above solution is optimal if for all  $i$

$$V_\beta(i,0) \leq V_\beta(i,1) + R_1$$

$$V_\beta(i,1) \leq V_\beta(i,0) + R_2 .$$

In the former case we have for all  $i$

$$A_1 + B_1 i \leq A_2 + B_2 i + K_2 G^i + R_1 ,$$

i.e.,

$$\psi_1 \geq \max_i H\left(1 - \frac{(1-A)}{(1-AG)} G^i\right) = H.$$

In the latter case we have for all  $i$

$$A_2 + B_2 i + K_2 G^i \geq A_1 + B_1 i + R_2$$

i.e.,

$$\psi_1 \leq \min_i H(1 - \frac{(1-A)}{(1-AG)} G^i) + R_1 + R_2 = HA \frac{(1-G)}{(1-AG)} + R_1 + R_2.$$

If  $\psi_1 < h$ , there will be an  $\hat{n}$  such that

$$\psi_1 < H(1 - \frac{(1-A)}{(1-AG)} G^{\hat{i}})$$

for all  $i \geq \hat{n}$ . From the policy improvement algorithm the total discounted cost will be reduced by making all  $i \geq \hat{n}$  service integers.

For  $\psi_1 < H$  we consider the solution of the transcendental equation

$$\psi_1 = H(1-G^x)$$

giving

$$x = \frac{\log(\frac{H-\psi_1}{H})}{\log G} . \quad (2.8)$$

We also set  $n_o = \langle x \rangle$ , where  $\langle x \rangle$  is the smallest integer greater than or equal to  $x$ .

**THEOREM 2.13.**

If  $\psi_1 < H$  and  $R_1 + R_2 \geq \frac{H(1-A)}{(1-AG)}$ , then that policy which makes all  $i \geq n_o$  service integers and all non-negative  $i < n_o$  indifference integers is a  $\beta$ -optimal improvement policy.

Proof. In this case we solve (2.5) for all  $i$ , (2.4) for all  $i$  with  $0 \leq i < n_0$ , and (2.6) for all  $i \geq n_0$ . The solution requires that

$$v_\beta(n_0, 0) = v_\beta(n_0, 1) + R_1,$$

which gives

$$K_1 = A^{n_0} (\psi_1 - H) \left( 1 - \frac{(1-A)}{(1-AG)} G^{n_0} \right)$$

and

$$K_2 = H \frac{(1-A)}{(1-AG)} .$$

For this solution to be optimal we must have

$$v_\beta(i, 0) - \frac{\lambda}{\lambda+\beta} v_\beta(i+1, 0) \leq \frac{h_i + r_1}{\lambda + \beta} . \quad i \geq n_0$$

This is equivalent to

$$A_2 + B_2 i + K_2 G^i + R_1 - \frac{\lambda}{\lambda+\beta} (A_2 + B_2 i + B_2 + K_2 G^{i+1} + R_1) \leq \frac{h_i + r_1}{\lambda + \beta},$$

which reduces to

$$\psi_1 \leq H(1-G^i). \quad i \geq n_0$$

Since  $H(1-G^i)$  is increasing in  $i$ , and by definition this inequality is satisfied for  $i = n_0$ , it is satisfied for all  $i \geq n_0$ .

Also for  $i \geq n_0$  we must have

$$v_\beta(i,1) \leq v_\beta(i,0) + R_2 = v(i,1) + R_1 + R_2,$$

or equivalently

$$0 \leq R_1 + R_2 \quad i \geq n_0$$

which is satisfied.

For  $i < n_0$  we must have

$$v_\beta(i,0) \leq v_\beta(i,1) + R_1,$$

i.e.,

$$A_1 + B_1 i + K_1 A^{-i} \leq A_2 + B_2 i + K_2 G^i + R_1.$$

Substitution for  $K_1$  and  $K_2$  yields the equivalent condition

$$\psi_1 \geq H - HG^i \frac{(1-A)(1-(AG)^{n_0-i})}{(1-AG)(1-A^{n_0-i})} \quad (2.9)$$

we add and subtract  $HG^{n_0-1}$  to the right hand side of (2.9) above. An equivalent condition is

$$\begin{aligned}
\psi_1 &\geq H(1 - G^{n_0^{-1}}) - HG^i \frac{(1-A)}{(1-A)^{n_0^{-i}}} [1 + AG + \dots + (AG)^{n_0^{-i-1}} \\
&\quad - G^{n_0^{-i-1}} (1 + A + \dots + A^{n_0^{-i-1}})] \\
&= H(1 - G^{n_0^{-1}}) - HG^i \frac{(1-A)}{(1-A)^{n_0^{-i}}} [1 - G^{n_0^{-i-1}} + \\
&\quad + AG(1 - G^{n_0^{-i-2}}) + \dots + (AG)^{n_0^{-i-2}} (1-G)],
\end{aligned}$$

where the second expression on the right is non-negative for all  $i \leq n_0^{-1}$ .

Since  $\psi_1 > H(1 - G^{n_0^{-1}})$ , the desired result follows.

Also for  $i < n_0$  we must have

$$v_\beta(i,1) \leq v_\beta(i,0) + R_2,$$

i.e.,

$$(\psi_1 - H)(1 - A^{n_0^{-i}}) \leq R_1 + R_2 - HG^i \frac{(1-A)}{(1-AG)} (1 - (AG)^{n_0^{-i}}).$$

Since

$$R_1 + R_2 \geq H \frac{(1-A)}{(1-AG)} \geq H \frac{(1-A)}{(1-AG)} G^i ,$$

the above inequality holds if

$$(\psi_1 - H)(1 - A^{n_0^{-i}}) \leq H \frac{(1-A)}{(1-AG)} (AG)^{n_0^{-i}} G^i ,$$

i.e., if

$$\psi_1 \leq H + H \frac{(1-A)(AG)^{n_0-i} G^i}{(1-AG)(1-A^{n_0-i})}.$$

Since by assumption  $\psi_1 \leq H$ , the above inequality is satisfied if  $i < n_0$ .

Corollary 2.14. If  $\psi_1 \leq H$  and  $\psi_1 > 0$ ,  $n_0 > 0$ .

Proof. The proof is immediate from the solution for  $\psi_1$  in (2.9) and the fact that  $n_0 \geq x > 0$ .

If  $R_1 + R_2 < H(1-A)/(1-AG)$ , it may be possible to introduce idle integers and obtain a policy improvement. Toward this end we consider the following transcendental equation in  $y$ :

$$(H - \psi_1)[(1-A)(1-(AG)^y - G^y(1-AG)(1-A^y))] = (R_1 + R_2)(1-AG)G^y \quad (2.10)$$

and define

$$n_1 = \langle y \rangle,$$

providing (2.10) has a solution. A sufficient condition that (2.10) has a solution is contained in the following lemma.

Lemma 2.15. If  $\psi_1 < H$ , there exists an  $n_1$  such that

$$(H - \psi_1) [(1-A)(1-(AG)^i) - G^i(1-AG)(1-A^i)] \geq (R_1 + R_2)(1-AG)G^i \quad i \geq n_1$$

$$(H - \psi_1) [(1-A)(1-(AG)^i) - G^i(1-AG)(1-A^i)] < (R_1 + R_2)(1-AG)G^i \quad 0 \leq i < n_1.$$

Proof. Since

$$(1-A)(1-(AG)^i) - G^i(1-AG)(1-A^i)$$

$$= (1-A)(1-AG)[1 - G^i + AG(1-G^{i-1}) + \dots + (AG)^{i-1}(1-G)] \geq 0$$

and  $H - \psi_1 > 0$ , both sides of the inequality above are non-negative. The right hand side of (2.10) is strictly decreasing in  $i$ . Moreover,

$$\begin{aligned} & (1-A)(1-(AG)^{i+1}) - G^{i+1}(1-AG)(1-A^{i+1}) - (1-A)(1-(AG)^i) - G^i(1-AG)(1-A^i) \\ & = G^i(1-AG)(1-G)(1-A^{i+1}) > 0, \end{aligned}$$

so the left side of (2.10) is strictly increasing in  $i$ . Since the left side of (2.10) is 0 for  $y = 0$  and the right side of (2.10) is  $> 0$  for  $y = 0$ , the result follows.

We now prove a somewhat stronger form of Theorem 2.13.

#### THEOREM 2.16.

If  $\psi_1 < H$  and  $n_0 < n_1$ , then that policy which makes all  $i \geq n_0$  service integers and all non-negative  $i < n_0$  indifference integers is a  $\beta$ -optimal improvement policy.

Proof. The proof is exactly the same as in Theorem 2.13 until we check if

$$v_\beta(i,1) \leq v_\beta(i,0) + R_2,$$

i.e., for  $i < n_0$

$$(\psi_1 - H)(1 - A^{n_0^{-i}}) \leq R_1 + R_2 - HG \frac{i(1-A)}{(1-AG)} (1 - A^{n_0^{-i}} G^{n_0^{-i}}) \quad (2.11)$$

From Lemma 2.15 with  $n_0 < n_1$

$$\begin{aligned} & (H - \psi_1) [(1-A)(1-(AG))^{n_0^{-i}}] - G^{n_0^{-i}} (1-AG)(1 - A^{n_0^{-i}}) \\ & < (R_1 + R_2)(1-AG)G^{n_0^{-i}}. \end{aligned} \quad (2.12)$$

Inequality (2.11) may be rewritten as

$$-G^{n_0^{-i}} (H - \psi_1)(1 - A^{n_0^{-i}})(1-AG) \leq G^{n_0^{-i}} (R_1 + R_2)(1-AG) - G^{n_0^{-i}} H (1-A)(1-(AG))^{n_0^{-i}}.$$

Adding and subtracting  $(H - \psi_1)(1-A)(1-(AG))^{n_0^{-i}}$  to the left hand side of

(2.11) yields

$$\begin{aligned} & (H - \psi_1) [(1-A)(1-(AG))^{n_0^{-i}}] - G^{n_0^{-i}} (1-AG)(1 - A^{n_0^{-i}}) - (H - \psi_1)(1-A)(1-(AG))^{n_0^{-i}} \\ & \leq G^{n_0^{-i}} (R_1 + R_2)(1-AG) - G^{n_0^{-i}} H (1-A)(1-(AG))^{n_0^{-i}}. \end{aligned}$$

From (2.12) above this inequality will be satisfied if

$$(H - \psi_1) \geq G^{n_0^{-i}} H$$

which is true by definition of  $n_0$ .

We now consider that policy which makes  $i = 0$  an idle integer,  $i \geq n$  service integers, and all positive  $i < n$  indifference integers.

THEOREM 2.17.

If  $\psi_1 < H$  and  $n_1 < n_0$ , then that policy which makes all  $i \geq n_1$  service integers, all positive  $i < n_1$  indifference integers and  $i = 0$  an idle integer is a  $\beta$ -optimal improvement policy.

Proof. In this case we solve (2.4) for all  $i$  such that  $0 \leq i < n_1$ , (2.5) for all  $i \geq 1$ , (2.6) for all  $i \geq n_1$ , and (2.7) for  $i = 0$ . The solution is of the form

$$v_\beta(i,0) = A_1 + B_1 i + K_1 A^{-i}$$

$$v_\beta(i,1) = A_2 + B_2 i + K_2 G^i,$$

where  $K_1$  and  $K_2$  are determined by the boundary conditions

$$v_\beta(n_1,0) = v(n_1,1) + R_1$$

$$v_\beta(0,1) = v(0,0) + R_2,$$

which implies

$$K_1 - K_2 (AG)^{n_1} = A^{n_1} (\psi_1 - H)$$

$$-K_1 + K_2 = R_1 + R_2 - (\psi_1 - H),$$

i.e.,

$$K_2 = \frac{(R_1 + R_2) + (H - \psi_1)(1 - A^{-1})}{1 - (AG)^{n_1}}$$

$$K_1 = \frac{(AG)^{n_1}(R_1 + R_2) - A^{n_1}(1 - G^{n_1})(H - \psi_1)}{1 - (AG)^{n_1}}.$$

For optimality we must have for  $i \geq n_1$

$$v_\beta(i, 0) \leq \frac{\lambda}{\lambda + \beta} v_\beta(i+1, 0) + \frac{h_i + r_1}{\lambda + \beta}$$

which reduces to the condition

$$(H - \psi_1) \geq \frac{[(R_1 + R_2) + (H - \psi_1)(1 - A^{-1})]G^i(1 - AG)}{(1 - A)(1 - (AG)^{n_1})}.$$

Clearly the right hand side is decreasing in  $i$ , and from Lemma 2.15 the inequality is satisfied for  $i = n_1$ , so that it holds for all  $i \geq n_1$ .

Also for  $i \geq n_1$  we must have

$$v_\beta(i, 1) \leq v_\beta(i, 0) + R_2.$$

Since

$$v_\beta(i,0) = v_\beta(i,1) + R_1,$$

this condition is satisfied.

For  $i < n_1$  we check if

$$v_\beta(i,0) \leq v_\beta(i,1) + R_1,$$

which after some algebraic manipulation results in the condition

$$\begin{aligned} (H - \psi_1) [(1-A)^{n_1-i} (1-(AG))^{n_1}] &= G^i (1-(AG))^{n_1-i} (1-A)^{n_1} \\ &\leq G^i (R_1 + R_2) (1-(AG))^{n_1-i}. \end{aligned}$$

It may be shown that for  $0 \leq i \leq n_1$ , the coefficient of  $H - \psi_1$  is non-negative. Recalling that

$$(H - \psi_1) < \frac{(R_1 + R_2) G^{n_1-1} (1-AG)}{(1-A)(1-(AG))^{n_1-1} - G^{n_1-1} (1-AG)(1-A)^{n_1-1}},$$

the condition for optimality will be satisfied if

$$\begin{aligned} &\frac{G^{n_1-1} (1-AG)}{(1-A)(1-(AG))^{n_1-1} - G^{n_1-1} (1-AG)(1-A)^{n_1-1}} \\ &\leq \frac{G^i (1-(AG))^{n_1-i}}{(1-A)^{n_1-i} (1-(AG))^{n_1} - G^i (1-(AG))^{n_1-i} (1-A)^{n_1}} \end{aligned}$$

which results in

$$\begin{aligned} & G^i (1-(AG)^{n_1-i}) [(1-A)(1-(AG)^{n_1-1}) - G^{n_1-1} (1-AG)(1-A^{n_1-1}) + G^{n_1-1} (1-AG)(1-A^{n_1})] \\ & \geq G^{n_1-1} (1-AG)(1-A^{n_1-1})(1-(AG)^{n_1}). \end{aligned}$$

Since

$$(1-A)(1-(AG)^{n_1-1}) - G^{n_1-1} (1-AG)(1-A^{n_1-1}) = (1-A)(1-(AG)^{n_1}),$$

the condition for optimality will be satisfied if

$$G^i (1-(AG)^{n_1-i})(1-A) \geq G^{n_1-1} (1-AG)(1-A^{n_1-i}),$$

i.e., if

$$G^i (1-AG)(1-A) [1-G^{n_1-1-i} + (AG)(1-G^{n_1-i-2}) + \dots + (AG)^{n_1-i-2} (1-G)] \geq 0.$$

Since this inequality holds for  $0 \leq i \leq n_1-1$ , the result follows.

For  $0 < i < n_1$  we check if

$$v_\beta(i,1) \leq v_\beta(i,0) + R_2,$$

which reduces to the equivalent condition

$$(H - \psi_1) [(1-A)^{n_1-i} (1-(AG)^i) - G^i (1-(AG))^{n_1-i} (1-A^i)] \\ \geq -(R_1 + R_2) [1 - G^i + A^{n_1-i} G^{n_1-i} (1-A^i)].$$

One may show that

$$(1-A)^{n_1-i} (1-(AG)^i) - G^i (1-(AG))^{n_1-i} (1-A^i) \\ = (1-A)(1-AG)[1-G^i + AG(1-G^{i-1}) + \dots + A^{i-1}G^{i-1}(1-G) \\ + A(1-G^{i+1}) + A^2G(1-G^i) + \dots + A^iG^{i-1}(1-G^2) \\ \vdots \\ + A^{n_1-i-1}(1-G^{n_1-i}) + A^{n_1-i}G^{n_1-i-2}(1-G^{n_1-i}) + \dots + A^{n_1-2}G^{i-1}(1-G^{n_1-i})].$$

It follows that the left hand side of the above inequality is non-negative and the associated condition for optimality is satisfied.

For  $i = 0$  we check if

$$v_\beta(0,1) \leq \frac{\lambda}{\lambda+\beta} v_\beta(1,1) + \frac{r_2}{\lambda+\beta}$$

which is equivalent to

$$(H - \psi_1) (1-A)^{n_1} (1-AG) - (1-A)(1-(AG))^{n_1} H \leq -(R_1 + R_2)(1-AG),$$

i.e.,

$$-G^{n_1} (H - \psi_1) (1 - A^{n_1}) (1 - AG) + G^{n_1} (1 - A) (1 - (AG)^{n_1}) H$$

$$+ (H - \psi_1) (1 - A) (1 - (AG)^{n_1}) - H (1 - A) (1 - (AG)^{n_1})$$

$$\geq G^{n_1} (1 - AG) (R_1 + R_2) - \psi_1 (1 - A) (1 - (AG)^{n_1}).$$

From Lemma 2.15 and the definition of  $n_1$ , the condition for optimality will be satisfied if

$$H(1 - G^{n_1}) \leq \psi_1.$$

Since  $n_1 < n_0$ , this condition is satisfied.

#### THEOREM 2.18.

If  $n^* = n_0 = n_1$ ,  $\psi_1 < H$  and

$$(H - \psi_1) (1 - A^{n^*}) (1 - AG) - H (1 - A) (1 - (AG)^{n^*}) \geq -(R_1 + R_2) (1 - AG),$$

the policy of Theorem 2.16 is a  $\beta$ -optimal improvement policy. If  $n^* = n_0 = n_1$ ,  $\psi_1 < H$  and

$$(H - \psi_1) (1 - A^{n^*}) (1 - AG) - H (1 - A) (1 - (AG)^{n^*}) \leq -(R_1 + R_2) (1 - AG),$$

the policy of Theorem 2.17 is a  $\beta$ -optimal improvement policy.

Proof. In the proof of Theorem 2.16 we note that the argument for optimality goes through for  $n_0 \leq n_1$  providing we restrict our attention to  $i > 0$ .

For  $i = 0$  we must have

$$(H - \psi_1)(1 - A^{n_0})(1 - AG) - H(1 - A)(1 - (AG)^{n_0}) \geq -(R_1 + R_2)(1 - AG).$$

In the proof of Theorem 2.17 we have the condition

$$(H - \psi_1)(1 - A^{n_1})(1 - AG) - H(1 - A)(1 - (AG)^{n_1}) \leq -(R_1 + R_2)(1 - AG).$$

With  $n^* = n_0 = n_1$  at least one of these results must hold.

We now prove two additional theorems which allow a complete characterization of all optimal policies.

**THEOREM 2.19.**

If

$$\max\left\{\frac{HA(1-G)}{1-AG} + R_1 + R_2, H\right\} \leq \psi_1 \leq H + R_1 + R_2,$$

then that policy for which  $i = 0$  is an idle integer and all  $i \geq 1$  are indifference integers is a  $\beta$ -optimal improvement policy.

Proof. In this case

$$V_\beta(i, 0) = A_1 + B_1 i$$

$$V_\beta(i, 1) = A_2 + B_2 i + K_2 G^i,$$

where  $K_2$  is determined from

$$v_\beta(0,1) = v_\beta(0,0) + R_2,$$

so

$$K_2 = H - \psi_1 + R_1 + R_2.$$

For optimality we must have for  $i \geq 0$

$$v_\beta(i,0) \leq v_\beta(i,1) + R_1,$$

i.e.,

$$\psi_1 - H \geq -[H + R_1 + R_2 - \psi_1]G^i. \quad (2.13)$$

Since

$$H \leq \psi_1 \leq H + R_1 + R_2,$$

the left side of (2.13) is non-negative and the right hand side of (2.13) is non-positive.

For  $i \geq 0$  we must have

$$v_\beta(i,1) \leq v_\beta(i,0) + R_2$$

i.e.,

$$\psi_1 - H \leq [\psi_1 - H - (R_1 + R_2)]G^i + R_1 + R_2.$$

This inequality is immediately satisfied for  $i = 0$ , and for  $i > 0$  we have

$$(\psi_1 - H)(1-G^i) \leq (R_1 + R_2)(1-G^i),$$

i.e.,

$$\psi_1 \leq H + R_1 + R_2,$$

which holds by assumption.

Finally we must check if

$$v_\beta(0,1) \leq \frac{\lambda}{\lambda+\beta} v_\beta(1,1) + \frac{r_2}{\lambda+\beta},$$

i.e.,

$$\psi_1 \geq \frac{HA(1-G)}{(1-AG)} + R_1 + R_2,$$

which holds by assumption.

#### THEOREM 2.20.

If  $\psi_1 \geq H + R_1 + R_2$ , then that policy for which all  $i \geq 0$  are idle integers is a  $\beta$ -optimal improvement policy.

Proof. In this case

$$v_\beta(i,0) = A_1 + B_1 i,$$

$$v_\beta(i,1) = A_1 + B_1 i + R_2.$$

We must check if

$$v_\beta(i,0) \leq v_\beta(i,1) + R_1 = v_\beta(i,0) + R_1 + R_2$$

which is satisfied since

$$0 \leq R_1 + R_2.$$

Also for  $i \geq 1$  we must have

$$\begin{aligned} v_\beta(i,1) &= \sum_{k=0}^{\infty} v_\beta(i+k-1,1) \int \frac{(\lambda t)^k}{k!} e^{-(\lambda+\beta)t} dB(t) \\ &\leq (h_i + r_2) \frac{(1-\tilde{B}(\beta))}{\beta} + \frac{\lambda h}{\beta^2} (1-\tilde{B}(\beta) - \beta \int t e^{-\beta t} dB(t)). \end{aligned}$$

Substitution for  $v_\beta(i,1)$  yields

$$\begin{aligned} (A_1 + B_1 i + R_2)(1-\tilde{B}(\beta)) + B_1 (\tilde{B}(\beta) - \lambda \int t e^{-\beta t} dB(t)) \\ \leq (h_i + r_2) \frac{(1-\tilde{B}(\beta))}{\beta} + \frac{\lambda h}{\beta^2} (1-\tilde{B}(\beta) - \beta \int t e^{-\beta t} dB(t)). \end{aligned}$$

Reduction of terms yields

$$\frac{(r_2 - r_1)}{\beta} \geq H + R_2,$$

i.e.,

$$\psi_1 \geq H + R_1 + R_2,$$

which is satisfied by assumption.

For  $i = 0$  we check if

$$v_\beta(0,1) - \frac{\lambda}{\lambda+\beta} v_\beta(1,1) \leq \frac{r_2}{\lambda+\beta},$$

i.e.,

$$\frac{(r_2 - r_1)}{\beta} \geq R_2,$$

i.e.,

$$\psi_1 \geq R_1 + R_2,$$

which is satisfied since  $H \geq 0$ .

We now prove a theorem which allows us to replace the expression  $\beta$ -optimal improvement policy by  $\beta$ -optimal policy in all the preceding theorems.

#### THEOREM 2.21.

All  $\beta$ -optimal improvement policies of Theorems 2.12, 2.13, 2.16, 2.17, 2.18, 2.19, and 2.20 are  $\beta$ -optimal policies.

Proof. Since  $0 < G < 1$ ,  $v_\beta(i,j) \leq K_i$  for some  $K \geq 0$ , all  $i \geq 0$  and  $j = 0$  or 1 in all theorems in question. It follows from Reed [7] Theorem 3.17 that the implied policies are optimal if

$$\mu^{-1} = \int t dG(t) < \infty.$$

Since we have assumed throughout this investigation that  $\lambda > 0$  and  $\rho = \lambda\mu^{-1} < 1$ , the result follows.

At this point we define three basically different policies.

Definition: A policy of type A is determined by an integer  $n_0$  such that all  $i \geq n_0$  are service integers and all non-negative  $i < n_0$  are indifference integers. For  $n_0 = \infty$ , all integers are indifference integers.

Definition: A policy of type B is determined by an integer  $n_1$  such that all  $i \geq n_1$  are service integers, all positive  $i < n_1$  are indifference integers and 0 is an idle integer. If  $n_1 = \infty$ , all  $i \geq 1$  are indifference integers and  $i = 0$  is an idle integer.

Definition: A policy is of type C if all  $i \geq 0$  are idle integers.

We now give a theorem which completely characterizes all possible optimal policies.

THEOREM 2.22. A stationary optimal policy exists for any combination of queueing system parameters, and the appropriate policy is given in the following table:

Conditions	Optimal Policy
$\psi_1 \geq H + R_1 + R_2$	C
$\max\left\{\frac{HA(1-G)}{(1-AG)} + R_1 + R_2, H\right\} \leq \psi_1 \leq H + R_1 + R_2$	B ( $n_1 = \infty$ )
$H \leq \psi_1 \leq \frac{HA(1-G)}{(1-AG)} + R_1 + R_2$	A ( $n_0 = \infty$ )
$\psi_1 < H$ and $n_0 < n_1$ , or $\psi_1 < H$ , $n_0 = n_1$ , and $(H - \psi_1)^{n_0} (1-A)(1-AG) - H(1-A)(1-(AG))^{n_0}$ $\geq -(R_1 + R_2)(1-AG)$	A ( $n_0 < \infty$ )
$\psi_1 < H$ and $n_1 < n_0$ , or $\psi_1 < H$ , $n_1 = n_0$ , and $(H - \psi_1)^{n_1} (1-A)(1-AG) - H(1-A)(1-(AG))^{n_1}$ $\leq -(R_1 + R_2)(1-AG)$	B ( $n_1 < \infty$ )

For  $\psi_1 \geq H$  the results above are identical to Blackburn [2] Chapter 2 Lemmas 6 and 7. For  $\psi_1 < H$  the above results provide a characterization of optimal policies different from Balckburn's.

It should be noted that we have derived closed form expressions for all expected discounted costs given any starting state. In particular for type A policies

$$v_\beta(0,0) = A_1 + K_1$$

where  $A_1$  and  $K_1$  are given in the proof of Theorem 2.13.

For  $n_0 = k$ ,

$$v_\beta(0,0) = \frac{r_1}{\beta} + \frac{\lambda h}{\beta^2} + A^k (\psi_1 - H) + H(1-A)(AG)^k / (1-AG)$$

$$= C_{on}(k)$$

where  $C_{on}(k)$  is Eq. (3) in Bell [1]. Similarly, for type B policies with  $n_1 = k$  and results in the proof of Theorem 2.17

$$v_\beta(0,0) = \frac{r_1}{\beta} + \frac{\lambda h}{\beta^2} + A^k (\psi_1 - H) + (AG)^k \frac{\{(R_1 + R_2) + (H - \psi_1)(1 - A^k)\}}{1 - (AG)^k}$$

$$= C(k),$$

where  $C(k)$  is Eq. (2) in Bell [1].

### 3. Optimal Control of a Bulk Queueing System

Consider a single-server queueing system that is controlled by turning the server on and performing a bulk service, after which he is turned off. Customers arrive according to a Poisson process with rate  $\lambda > 0$ . Once bulk service begins, all customers in the queue at the time of service initiation are served. Bulk service times are non-negative, independent random variables with common distribution function  $B$ . It should be stressed that

B. in no way depends on the number of customers served in bulk, and all customers beginning service together are finished together. Customers arriving while the service facility is busy from a new queue; consequently, a bulk service may end with a queue of any length.

We shall also assume that

$$\mu^{-1} = \int t dB(t) < \infty$$

and

$$\mu_2 = \int t^2 dB(t) < \infty.$$

Decisions are made at the time of service completions, or with the arrival of a customer if the server is not busy. Since it is assumed that the server is shut down (at least momentarily) at the completion of service, the decision at all permissible decision times is whether to remain idle or provide service for customers currently in the queue. We let  $k = 0$  imply the decision to remain idle and we let  $k = 1$  imply the decision to provide service.

The state of the system at a decision time is  $i$ , the number of people in the queue immediately after an arrival (when the server is idle) or immediately after completion of a bulk service. As in Section 2,  $i$  is right continuous with respect to the time parameter. The only exception is that  $i$  is discontinuous at decision points for which  $k = 1$  when  $\mu^{-1} = 0$ .

There is a fixed charge  $R$  of providing the bulk service, independent of the number in the queue. There is a holding cost of  $h$  per unit time for each customer waiting for service. We shall make two alternative assumptions with regard to holding costs for customers during the service period.

Under one assumption holding costs for a customer end once he enters service; under the other assumption there is a holding cost of  $h$  per unit time for each customer while being served. If the former of these assumptions holds, we shall be concerned with problem 1. If the second of these alternative assumptions holds, we shall be concerned with problem 2. We seek a policy which minimizes the expected average cost per unit time for both problems 1 and 2.

### 3.1. Existence of a Stationary Optimal Policy

We proceed to verify assumptions A(1), A(2), and A(3) of Reed [7] Section 4.1 for the existence of a stationary optimal policy.

**LEMMA 3.1.** If  $\mu_2 = \int t^2 dB(t) < \infty$ , A(1) is satisfied.

Proof. If  $\mu^{-1} > 0$ , we consider that stationary policy  $f_0$  which continually provides service. It follows from Reed [7] and Appendix A that the difference equations associated with this policy are

$$v_i - \sum v_k \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \phi_{f_0} \mu^{-1} = s_0 + s_1 i \quad (3.1)$$

where

$$s_0 = R + \frac{\lambda h \mu_2}{2}$$

$$s_1 = 0$$

for problem 1 and

$$S_0 = R + \frac{\lambda h \mu_2}{2}$$

$$S_1 = h \mu^{-1}$$

for problem 2.

Setting

$$v_i = Bi$$

and substituting in (3.1) yields

$$B = S_1$$

$$\phi_{f_0} = \lambda S_1 + \frac{S_0}{\mu^{-1}} .$$

For problem 1

$$\phi_{f_0} = \frac{R + \frac{\lambda h \mu_2}{2}}{\mu^{-1}} .$$

For problem 2

$$\phi_{f_0} = \lambda \mu^{-1} h + \frac{R + \frac{\lambda h \mu_2}{2}}{\mu^{-1}} .$$

It is interesting to note that for problem 2 there is an optimal expected service time

$$\hat{\mu}^{-1} = \left[ \frac{R + \frac{\lambda h \mu_2}{2}}{\lambda h} \right]^{\frac{1}{2}}$$

for which

$$\hat{\phi}_{f_0} = 2 \left[ \lambda h (R + \frac{\lambda \mu^2}{2}) \right]^{1/2}.$$

If  $\mu^{-1} = 0$ , we consider that policy,  $f_1$ , which always waits for the arrival of a single customer before providing instantaneous service. In this case it is easily shown that

$$\phi_{f_1} = \lambda k.$$

LEMMA 3.2. Assumption A(2) holds.

Proof. If  $n$  customers arrive during time  $t$ , then it is a property of the Poisson distribution that the total cost associated with holding these customers in  $hnt/2$ , where  $E(t) = n/\lambda$ . Now let  $\pi$  be any policy for which the probability that  $n$  customers are allowed to arrive between the conclusion of one service and the start of the next service is given by  $P_n$  for  $n = 0, 1, \dots$ . If service is never provided, we set  $P_\infty = 1$ . Since  $R \geq 0$ , it can only add to the average cost per unit time. Neglecting this cost and all holding costs for customers who arrive while a service is in progress,

$$\begin{aligned} \phi_\pi &> \frac{\sum P_n \frac{hn}{2} \frac{n}{\lambda}}{\sum P_n \frac{n}{\lambda} + \mu^{-1}} \\ &> \frac{\sum_{n=0}^{N-1} P_n \frac{hn^2}{2\lambda} + \frac{hN}{2} \sum_{n=N}^{\infty} P_n \frac{n}{\lambda}}{\sum_{n=0}^{N-1} \frac{P_n n}{\lambda} + \sum_{n=N}^{\infty} \frac{P_n n}{\lambda} + \mu^{-1}} \end{aligned}$$

$$\begin{aligned}
 & \sum_{n=0}^{N-1} \frac{P_n n^2 h}{2\lambda} \\
 & + \frac{\sum_{n=N}^k \frac{P_n n}{\lambda}}{\mu^{-1} + \sum_{n=0}^{N-1} \frac{P_n n}{\lambda} + 1} \\
 = \lim_{k \rightarrow \infty} & \frac{\sum_{n=N}^k \frac{P_n n}{\lambda}}{\mu^{-1} + \sum_{n=0}^{N-1} \frac{P_n n}{\lambda} + 1}
 \end{aligned}$$

If  $\mu_\pi$ , the mean time between the end and start of service, is  $\infty$ , then

$$\lim_{k \rightarrow \infty} \sum_{n=N}^k \frac{P_n n}{\lambda} = \infty$$

and

$$\phi_\pi > hN/2$$

for all  $N$ , showing that  $\phi_\pi = \infty$ . Lemma 3.1 excludes such policies and hence  $\mu_\pi < \infty$ . Thus for any permissible policy  $\pi$  the mean recurrence time between bulk services is  $\mu_\pi + \mu^{-1} < \infty$ .

Each time bulk service is performed there is a probability,

$\int e^{-\lambda t} dB(t) > 0$ , that the state 0 will be entered. Thus, for all permissible policies 0 is a positive recurrent state with mean recurrence time of  $(\mu_\pi + \mu) / e^{-\lambda t} dB(t)$ .

LEMMA 3.3. Assumption A(3) holds.

Proof. If  $\mu^{-1} > 0$ ,

$$c_i(0) = \frac{hi}{\lambda},$$

$$c_i(1) = R + \frac{\lambda h \mu_2}{2} \quad \text{for problem 1,}$$

$$c_i(1) = R + \frac{\lambda h \mu_2}{2} + h \mu^{-1} i \quad \text{for problem 2,}$$

$$t_i(0) = \frac{1}{\lambda}, \text{ and}$$

$$t_i(1) = \mu^{-1}.$$

For problem 1

$$c_i(0) - \phi_{f_0}/\lambda = \frac{hi}{\lambda} - \frac{(R + \frac{\lambda h \mu_2}{2})}{\lambda \mu^{-1}},$$

$$c_i(1) - \phi_{f_0} \mu^{-1} = 0.$$

For problem 2

$$c_i(0) - \phi_{f_0}/\lambda = \frac{hi}{\lambda} - \mu^{-1} h - \frac{R + \frac{\lambda h \mu_2}{2}}{\lambda \mu^{-1}},$$

$$c_i(1) - \phi_{f_0} \mu^{-1} = h \mu^{-1} i - \lambda \mu^{-2} h.$$

It follows from Reed [7] Theorem 4.3 that  $V$  is bounded below.

If  $\mu^{-1} = 0$ , trivial sequences with non-negative costs are eliminated by not allowing service in state 0. In this case Condition 1 is satisfied and Reed [7] Theorem 4.3 becomes applicable by setting  $\mu^{-1} = \mu_2 = 0$

in  $C_i(j)$  and  $t_i(j)$  and computing  $C_i(j) - \phi_{f_1} t_i(j)$ . We summarize these results in a theorem.

**THEOREM 3.4.**

If  $\mu_2 < \infty$ , there exists a stationary optimal policy for bulk queueing problems 1 and 2.

**3.2. Qualitative Attributes of an Optimal Policy**

We now investigate the attributes of  $V_{f^*}$ , the cost function associated with an optimal policy  $f^*$ . From Reed [7], Theorem 4.1, the optimal cost function must satisfy

$$V_{f^*}(i) = \min[V_{f^*}(i+1) + \frac{hi}{\lambda} - \frac{\phi_{f^*}}{\lambda}, \\ \sum_{k=0}^{\infty} V_{f^*}(k) / e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + s_0 + s_1 i - \phi_{f^*} \mu^{-1}].$$

**LEMMA 3.5.** The set of  $i$  for which

$$V_{f^*}(i) = \sum_k V_{f^*}(k) / e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + s_0 + s_1 i - \phi_{f^*} \mu^{-1}$$

is unbounded.

**Proof.** Suppose the contrary is true, so that there exists an  $n$  such that

$$V_{f^*}(i) = \frac{hi}{\lambda} - \frac{\phi_{f^*}}{\lambda} + V_{f^*}(i+1) \quad i = n, n+1, \dots$$

Now suppose that an upper bound  $N > n$  is imposed on the number of customers allowed in the queue. Let  $\phi_{f^*}(N)$  be the average cost per unit time for this bounded problem under the policy  $f^*$ , where  $\phi_{f^*}(N) \leq \phi_{f^*}$ . For  $i = N$  we have

$$v_{f^*}(N) = \frac{hN}{\lambda} - \frac{\phi_{f^*}(N)}{\lambda} + v_{f^*}(N),$$

so that

$$\phi_{f^*}(N) = hN.$$

In the limit,  $\phi_{f^*}(N) \rightarrow \infty$ , contradicting the assumption that  $\phi_{f^*} < \infty$ .

LEMMA 3.6. The set of  $i$  for which

$$v_{f^*}(i) = \frac{hi}{\lambda} - \frac{\phi_{f^*}}{\lambda} + v_{f^*}(i+1)$$

is bounded.

Proof. Assume the contrary is true. From Lemma 3.5 it follows there is an unbounded sequence  $\{i_j\}$  such that

$$v_{f^*}(i_j) = \frac{hi_j}{\lambda} - \frac{\phi_{f^*}}{\lambda} + v_{f^*}(i_j+1)$$

$$v_{f^*}(i_j) \leq s_0 + s_1 i_j - \phi_{f^*}^{-1} + \sum v_{f^*}(k) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t)$$

$$v_{f^*}(i_j + 1) = s_0 + s_1(i_j + 1) - \phi_{f^*}^{-1} + \sum v_{f^*}(k) \int e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t).$$

Now since  $\phi_f^* < \infty$ , there exists an  $i_j$  such that

$$v_{f^*}(i_j) - v_{f^*}(i_j+1) > 0.$$

Also

$$\begin{aligned} v_{f^*}(i_j) - v_{f^*}(i_j+1) &\leq -s_1 \\ &\leq 0, \end{aligned}$$

since  $s_1 \geq 0$ . This gives a contradiction.

Since the possible actions in each state are (1) perform service immediately or (2) hold customers for future service, determining a stationary optimal policy is nothing more than deciding if each state  $i$  is a "service state" or a "holding state". Lemma 3.5 merely says that service states are unbounded and Lemma 3.6 says that holding states are bounded. We may restrict our search of an optimal stationary policy to policies with this attribute. We summarize this result in a theorem.

#### THEOREM 3.7.

If  $0 < \lambda < \infty$  and  $\mu_2 < \infty$ , then a stationary optimal policy exists and is characterized by having its set of holding states bounded.

#### 3.3. Determination of an Optimal Policy

We now show that an optimal improvement policy exists which is a special case of the policies described in Theorem 3.7. Finally, we shall show that this optimal improvement policy is optimal. We shall consider the following policy:

Hold customers until  $n$  are in the queue and then start bulk service. If there are  $n$  or more customers in the new queue at the end of bulk service, begin a new bulk service. If there are less than  $n$  at the end of bulk service, wait until there are  $n$  and then begin bulk service. We shall refer to the above policy as a monotone policy with parameter  $n$ .

The difference equations associated with this policy are

$$v_i - v_{i+1} + \frac{\phi_n}{\lambda} = \frac{hi}{\lambda} \quad i = 0, 1, 2, \dots, n-1 \quad (3.2)$$

$$v_i - \sum_{k=0}^{\infty} v_k \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \mu^{-1} \phi_n = s_0 + s_1 i \quad i = n, n+1, \dots, \quad (3.3)$$

where  $s_1 \geq 0$ .

In this case we attempt to find a solution of the form

$$v_i = b_i + c_i^2 \quad i = 0, 1, 2, \dots, n-1, \text{ and}$$

$$v_i = K_n^* + B^* i \quad i = n, n+1, \dots .$$

Substituting in (3.3) for  $i = n, n+1, \dots$ , we have

$$\begin{aligned} K_n^* + B^* i - \sum_{k=0}^{n-1} (bk + ck^2) \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) \\ - \sum_{k=n}^{\infty} (K_n^* + B^* k) \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \mu^{-1} \phi_n = s_0 + s_1 i. \end{aligned}$$

Setting  $B^* = s_1$ , we have

$$K_n^* P(n-1, 0) + \phi_n(\mu^{-1} - P(n-2, 1)) = s_0 - \frac{\lambda h}{2} P(n-3, 2) + s_1 \lambda (\mu^{-1} - P(n-2, 1)) \quad (3.4)$$

where

$$P(n, r) = \int_0^\infty \sum_{k=0}^n e^{-\lambda t} \frac{\lambda t^k}{k!} t^r dB(t)$$

is obtained in the course of interchanging summation and integration, which is permissible from the properties of the exponential function and the fact that  $B$  is a distribution function.

Substitution for  $i = n-1$  in (3.2) yields

$$\frac{n\phi_n}{\lambda} - K_n^* = \frac{hn(n-1)}{2\lambda} + s_1 n. \quad (3.5)$$

Clearly, equations (3.4) and (3.5) provide two linear equations in  $\phi_n$  and  $K_n^*$  which provide a solution for the average cost per unit time for such a policy. We shall consider the form of improvement that can occur if the policy improvement algorithm is applied to a monotone policy with parameter  $n$ .

If there exists  $i \geq n$  such that

$$v_i - v_{i+1} + \frac{\phi_n}{\lambda} > \frac{hi}{\lambda},$$

then an improvement in policy will be obtained by starting bulk service only after  $n^* > n$  customers have arrived. If, on the other hand, there exists  $i < n$  such that

$$v_i - \sum_{k=0}^{\infty} v_k f e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \mu \phi_n > s_0 + s_1 i,$$

Then there will be a set  $S(n)$  such that an improvement in policy will be obtained by starting service for all  $i \in S(n)$  and all  $i \geq n$ .

For  $i \geq n$ ,  $v_i = K_n^* + s_1 i$ , so the first condition becomes  $n \leq i < (\phi_n - \lambda s_1)/h$ , and we define

$$W(n) = \{i: n \leq i < (\phi_n - \lambda s_1)/h\}.$$

Since

$$K_n^* + s_1 i - \sum_{k=0}^{\infty} v_k \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dB(t) + \mu^{-1} \phi_n = s_0 + s_1 i,$$

the second condition becomes

$$v_i > K_n^* + s_1 i \quad i < n.$$

This in turn is equivalent to

$$(\frac{\phi_n}{\lambda} + \frac{h}{2\lambda})i - \frac{h}{2\lambda} i^2 > K_n^* + s_1 i.$$

Substitution for  $K_n^*$  from (3.5) gives for  $n > i$

$$i > \frac{2(\phi_n - \lambda s_1)}{h} - (n-1),$$

and we define

$$S(n) = \{i: 2(\phi_n - \lambda s_1)/h - (n-1) < i \leq n-1\}.$$

If  $\emptyset$  is the empty set and  $S(n) = \emptyset$  and  $W(n) = \emptyset$ , the associated monotone policy cannot be improved by policy improvement and it will be a  $\alpha$ -optimal improvement policy. We now consider the effect of applying the

policy improvement algorithm to a monotone policy.

LEMMA 3.8. If  $S(n) \neq \emptyset$ , then  $W(n) = \emptyset$ . Thus, if  $W(n) \neq \emptyset$ , then  $S(n) = \emptyset$ .

Proof. Assume there exists  $i_1 \geq n$  such that

$$i_1 < \frac{(\phi_n - \lambda s_1)}{h},$$

and there exists  $i_0 \leq n-1$  such that

$$\frac{(\phi_n - \lambda s_1)}{h} < \frac{n+i_0-1}{2}.$$

Therefore,

$$\frac{\phi_n - \lambda s_1}{h} < \frac{n+i_0-1}{2} \leq n-1 < n \leq i_1 < \frac{\phi_n - \lambda s_1}{h},$$

a contradiction.

LEMMA 3.9. Policy improvement of a monotone policy leads to a monotone policy.

Proof. If  $W(n) \neq \emptyset$ , we define

$$n^* = \max\{i: n \leq i < (\phi_n - \lambda s_1)/h\} + 1.$$

the monotone policy with parameter  $n^* \geq n+1$  is a result of the policy improvement.

If  $S(n) \neq \Phi$ , then for all  $i \in S(n)$ , inequality (3.6) is satisfied.

We observe that if there exists  $i < n$  for which (3.6) is satisfied, then the inequality (3.6) will hold for  $i, i+1, \dots, n-1$ . In this case, if we define  $n^* = \min\{i: i \in S(n)\}$ , then the policy improvement algorithm leads to a monotone policy with parameter  $n^*$ .

LEMMA 3.10. If  $n_{j_i}$  is the parameter associated with the  $j_i^{\text{th}}$  iteration of the policy improvement algorithm and  $j_i$  is the iteration on which  $W(n_{j_i-1}) \neq \Phi$  for the  $i^{\text{th}}$  time, then  $W(n_{j_i}) = \Phi$  and  $n_{j_i} < n_{j_{i-1}} < \dots < n_{j_1}$ .

#### Proof.

Part (i)  $W(n_{j_i}) = \Phi$ .

The  $i^{\text{th}}$  iteration leads to

$$n_{j_i} = \max\{k: n_{j_i-1} \leq k < (\phi n_{j_i-1} - \lambda s_1)/h\} + 1,$$

from which it follows that

$$(\phi n_{j_i-1} - \lambda s_1)/h \leq n_{j_i}.$$

If on step  $j_i+1$ ,  $W(n_{j_i}) \neq \Phi$ , then there would exist an  $i_1$  such that

$$n_{j_i} \leq i_1 < (\phi n_{j_i} - \lambda s_1)/h,$$

which implies

$$\phi n_{j_i-1} < \phi n_{j_i},$$

a contradiction. Hence  $W(n_{j_i}) = \Phi$  and either  $S(n_{j_i}) = \Phi$  or  $S(n_{j_i}) \neq \Phi$ .

In the former case the policy improvement algorithm has converged. In the latter case the policy improvement algorithm goes on.

Part (ii)  $\{n_{j_i}\}$  is strictly decreasing in  $i$ :

If  $S(n_{j_i}) \neq \Phi$ , then on step  $j_i+1$  there exists  $i_0 < n_{j_i}$  such that

$$\frac{\phi n_{j_i} - \lambda S_1}{h} < \frac{n_{j_i} + i_0 - 1}{2} \leq n_{j_i-1},$$

so that  $n_{j_i} > n_{j_i+1}$ . If, moreover,  $S(n_{j_i+m}) \neq \Phi$  for  $m = 1, 2, \dots, r$ , we obtain in the same way a sequence  $n_{j_i} > n_{j_i+1} > \dots > n_{j_i+r}$ , stopping only when  $S(n_{j_i+r+1}) = \Phi$ . If  $W(n_{j_i+r+1}) = \Phi$ , the policy improvement algorithm has converged. If  $W(n_{j_i+r+1}) \neq \Phi$ , then

$$n_{j_{i+1}} = \max\{k: n_{j_{i+1}-1} \leq k < (\phi n_{j_{i+1}-1} - \lambda S_1)/h\} + 1.$$

Now

$$\frac{\phi n_{j_{i+1}-1} - \lambda S_1}{h} \leq \frac{\phi n_{j_i} - \lambda S_1}{h} < n_{j_i} - 1,$$

so

$$n_{j_{i+1}} < \frac{\phi n_{j_{i+1}-1} - \lambda S_1}{h} + 1 < n_{j_i} - 1 + 1 = n_{j_i},$$

and

$$n_{j_{i+1}} < n_{j_i}.$$

THEOREM 3.11.

The Policy Improvement Algorithm terminates with a monotone, 0-optimal improvement policy.

Proof. Application of the policy improvement algorithm gives rise to a sequence of the following form:

$$n_{j_0} > n_{j_0+1} > \dots > n_{j_1-1}$$

$$n_{j_1} > n_{j_1+1} > \dots > n_{j_2-1}$$

$$n_{j_2} > n_{j_2+1} > \dots > n_{j_3-1}$$

⋮

$$n_{j_k} > n_{j_k+1} > \dots > n_{j_{k+1}-1}$$

Since

$$n_{j_1} > n_{j_2} > \dots > n_{j_k},$$

this sequence must stop in a finite number of steps with some  $n^*$  with  $W(n^*) = S(n^*) = \Phi$ . Otherwise,  $n_j \rightarrow -\infty$  and clearly,  $n_j \geq 0$  for all  $j$ .  
The resulting policy from Lemma 3.9 is monotone with parameter  $n^*$ .

THEOREM 3.12.

The monotone, 0-optimal improvement policy of Theorem 3.11 is optimal.

Proof. We let  $f^*$  be that stationary monotone policy with parameter  $n^*$  obtained from convergence of the policy improvement algorithm. If  $s_1 = 0$ ,

$$v_{f^*}(i) = \left(\frac{\phi_n}{\lambda} + \frac{h}{2\lambda}\right)i - \frac{h}{2\lambda} i^2 \quad i = 0, 1, 2, \dots, n^*-1$$

$$v_{f^*}(i) = K_{n^*}^* \quad i = n^*, n^*+1, \dots .$$

Clearly the  $v_{f^*}(i)$  are bounded and from Reed [7], Corollary 4.9 the resulting stationary policy is optimal.

If  $s_1 > 0$ ,

$$v_{f^*}(i) = \left(\frac{\phi_n}{\lambda} + \frac{h}{2\lambda}\right)i - \frac{h}{2\lambda} i^2 \quad i = 0, 1, \dots, n^*-1$$

$$v_{f^*}(i) = K_{n^*}^* + s_1 i \quad i = n^*, n^*+1, \dots .$$

In this case the  $v_{f^*}(i)$  are unbounded and from Reed [7], Theorem 4.8 a sufficient condition for optimality is  $x(f)v_{f^*} < \infty$  for all  $f \in \mathcal{F}$ . Lemma 3.6 shows that the Markov matrix of all feasible stationary policies  $f$  will be of the form

$$P(f) = \begin{bmatrix} P_{00}(f) & P_{01}(f) & \dots & P_{0n}(f) & \dots \\ P_{10}(f) & P_{11}(f) & \dots & P_{1n}(f) & \dots \\ P_{N-10}(f) & P_{N1}(f) & \dots & P_{Nn}(f) & \dots \\ P_0 & P_1 & & P_n & \\ \vdots & & & \vdots & \\ P_0 & P_1 & & P_n & \\ \vdots & & & \vdots & \end{bmatrix}$$

where

$$P_i = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^i}{i!} dB(t).$$

It follows that

$$x_i(f) = \sum_{j=0}^{N-1} x_j(f) P_{ji}(f) + P_i \sum_{j=N}^{\infty} x_j(f) \quad i = 0, 1, 2, \dots$$

There are only two possibilities for  $P_{ji}(f)$  for  $j = 0, 1, 2, \dots, N-1$ .

Either  $P_{ji}(f) = P_i$  for all  $i$  or  $P_{ji}(f) = 0$  for  $i > N$ . Let

$$T = \{j: 0 \leq j \leq N-1 \text{ and } P_{ji}(f) = 0 \text{ for } i > N\}.$$

Now

$$x_i(f) = \sum_{j \in T} x_j(f) P_{ji}(f) + P_i \sum_{j \in \bar{T}} x_j(f)$$

where  $\bar{T}$  is the complement of  $T$  on  $\{0, 1, 2, \dots\}$ . Now

$$\begin{aligned} \sum_{i=0}^{\infty} x_i(f) v_{f^*}(i) &= \sum_{i=0}^N \sum_{j \in T} x_j(f) p_{ji}(f) v_{f^*}(i) + \sum_{i=0}^{\infty} v_{f^*}(i) \sum_{j \in T} x_j(f) \\ &= \sum_{i=0}^N \sum_{j \in T} x_j(f) p_{ji} v_{f^*}(i) + \sum_{j \in T} x_j(f) \sum_{i=0}^{\infty} v_{f^*}(i). \end{aligned}$$

We observe that  $v_{f^*}(i) \leq A + Bi$  with  $B > 0$  and

$$\sum_{i=0}^N \sum_{j \in T} x_j(f) p_{ji} v_{f^*}(i)$$

is clearly bounded. Also  $\sum_{j \in T} x_j(f)$  is bounded and  $\sum_{i=0}^{\infty} x_i(f) v_{f^*}(i) < \infty$

if and only if  $\sum_{i=0}^{\infty} p_i v_{f^*}(i)$  is bounded, and this expression is bounded

if and only if  $\sum_{i=0}^{\infty} i p_i < \infty$ . The generating function associated with  $p_i$

is

$$K(z) = \sum_{j=0}^{\infty} p_j z^j = \int_0^{\infty} e^{-\lambda t + \lambda t z} dB(t),$$

with mean

$$K'(1) = \int_0^{\infty} \lambda t dB(t) = \lambda \mu^{-1}.$$

We have assumed  $0 < \lambda < \infty$  and  $\mu^{-1} < \infty$ , so  $\lambda \mu^{-1} < \infty$  and the sufficient condition for optimality is satisfied.

#### 4. The M/M/1 Queue with Variable Service Rate

Consider an M/M/1 queue that is controlled by selecting one of two service rates. Customers arrive according to a Poisson process with rate  $\lambda > 0$ . Service times are non-negative, independent random variables with an exponential distribution with service rate  $\mu_1$  or  $\mu_2$ , with  $\mu_1 < \mu_2 < \infty$ . We shall assume that

$$\rho_2 = \frac{\lambda}{\mu_2} < 1.$$

There is a cost  $r_i$  per unit time of operating the queueing system with service rate  $\mu_i$ ,  $i = 1, 2$ . It is assumed  $r_1 < r_2$ . If  $i$  customers are in the queueing system, there is a holding cost of  $c_i^*$ , where  $c^* < c_0 < c_1 \dots$ ,  $c_i \rightarrow \infty$  as  $i \rightarrow \infty$  and  $c^* > -\infty$  is a lower bound on the holding cost rate.

Decisions are made at the time of service completion or the arrival of a customer in the system. We associate  $k = 1$  with the decision to use service rate  $\mu_1$  and we associate  $k = 2$  with the decision to use service rate  $\mu_2$ .

The state of the system is described by the pair  $(i, j)$  where  $i$  indicates the number of customers in the queue and  $j$  implies service rate  $\mu_j$  is in use. As in Section 2,  $i$  is right continuous and  $j$  is left continuous with respect to the time parameter.

The optimization criterion is minimum expected average cost per unit time.

##### 4.1. Existence of a Stationary Optimal Policy

We now proceed to verify assumptions A(1), A(2), and A(3) of Reed [6]

Section 4.1 for the existence of a stationary optimal policy. It should be noted that in the two preceding examples we made an explicit assumption about the second moment of the general service distribution, to assure that costs over a service period are finite. In this particular example we must assume that the  $c_i$  are such that there exists a policy  $\pi_0$  for which  $\phi_{\pi_0} < \infty$ .

We shall first prove a lemma giving a sufficient condition that A(1) be satisfied.

**LEMMA 4.1.** If  $\sum_{i=0}^{\infty} \rho_2^i c_i < \infty$ , A(1) is satisfied.

Proof. Let  $f_0$  be the stationary policy which always uses fast service.

From Reed [7] Section 2.2 we have

$$\phi_{f_0} = \frac{\sum_{i=1}^{\infty} x_i (c_i + r_2)/(\lambda + \mu_2) + x_0 (c_0 + r_2)/\lambda}{\sum_{i=1}^{\infty} x_i / (\lambda + \mu_2) + x_0 / \lambda}$$

where

$$x_i = \frac{\mu}{\lambda + \mu} x_{i+1} + \frac{\lambda}{\lambda + \mu} x_{i-1} \quad i = 1, 2, \dots$$

$$x_0 = \frac{\mu}{\lambda + \mu} x_1 .$$

It follows that

$$x_i = \rho^{i-1} (1 - \rho^2) / 2 \quad i = 1, 2, \dots$$

$$x_0 = (1 - \rho) / 2$$

and

$$\phi_{f_0} = r_2 + \sum_{i=0}^{\infty} (1-\rho_2)\rho_2^i c_i < \infty.$$

LEMMA 4.2. If A(1) holds, A(2) is satisfied.

Proof. A(1) implies the existence of a policy  $f_0$  for which  $\phi_{f_0} < \infty$ .

Now we must show there exists a state recurrent over all policies  $\pi$  for which  $\phi_\pi \leq \phi_{f_0}$ . We define

$$I_1 = \{(i,1) : c_i + r_1 \leq \phi_{f_0}\} \equiv \{(0,1) \dots (i_1,1)\}$$

$$I_2 = \{(i,2) : c_i + r_2 \leq \phi_{f_0}\} \equiv \{(0,2) \dots (i_2,2)\}.$$

Regardless of the number in the queue the probability of entering  $\overline{I_1 \cup I_2}$  while a service is being performed exceeds

$$P = \frac{\mu_1}{\mu_2} \left(\frac{\lambda}{\lambda+\mu_2}\right)^{i_1+1}.$$

Hence  $\overline{I_1 \cup I_2}$  is recurrent with expected transition time from  $I_1 \cup I_2$  less than

$$\frac{\mu_2}{\mu_1} \left(\frac{\lambda+\mu_2}{\lambda}\right)^{i_1+1}.$$

In going from  $\overline{I_1 \cup I_2}$  to  $I_1 \cup I_2$  one of the states  $\{(i_1,1), (i_1-1,1) \dots (i_2,1), (i_2,2)\}$ , must be entered. There are  $i_1 - i_2 + 2$  of these states. Since this finite set is recurrent, at least one of these states must be positive recurrent under any policy  $\pi$  for which  $\phi_\pi \leq \phi_{f_0}$ . Let  $s$  be

any of these states. Once  $s$  is entered the probability of bringing the queue to zero before the arrival of a single customer is at least  $(\mu_1^i / (\lambda + \mu_1))^{1/2}$ . The process is in either  $(0,1)$  or  $(0,2)$  when this occurs. Since  $r_1 < r_2$ , the average cost per unit time may always be reduced by making decision  $k = 1$  in  $(0,2)$ . This restriction on the class of permissible policies in no way eliminates a possible optimal policy and we may assume that whenever the queue is empty, the process is in state  $(0,1)$ . It follows that the event, "Enter  $s$  and reach  $(0,1)$  before the arrival of single customer," is positive recurrent for all permissible  $\pi$ . Clearly, the state  $(0,1)$  is positive recurrent for all  $\pi$ .

LEMMA 4.3. Assumption A(3) holds.

Proof. We use Reed [7] Theorem 4.3 and Appendix A. We have

$$c_{(i,1)}^{(1)} = \frac{c_i + r_1}{\lambda + \mu_1}$$

$$c_{(i,1)}^{(2)} = 0$$

$$c_{(i,2)}^{(2)} = \frac{c_i + r_2}{\lambda + \mu_2}$$

$$c_{(i,2)}^{(1)} = 0$$

$$t_{(i,1)}^{(1)} = \frac{1}{\lambda + \mu_1}$$

$$t_{(i,1)}^{(2)} = 0$$

$$t_{(i,2)}^{(2)} = \frac{1}{\lambda + \mu_2}$$

$$t_{(i,2)}^{(1)} = 0.$$

The exclusion of trivial sequences with non-negative costs leads to the verification of Condition 1 and justifies the inclusion of  $t_{(i,1)}(2) = t_{(i,2)}(1) = 0$ . Since  $c_i \rightarrow \infty$  with  $c_i < c_{i+1}$  and  $c_i \geq c^*$  the set  $S$  of Reed [7] Theorem 4.3 is finite. It follows that  $V$  is bounded below.

We summarize these results in a theorem.

#### THEOREM 4.4.

If  $\sum \rho_2^i c_i < \infty$ , then there exists a stationary optimal policy.

Crabill [3] Chapter VI investigates the  $K$  service rate problem. For this problem we have service rates  $\mu_1 < \mu_2 < \dots < \mu_K$  with corresponding operating costs  $r_1 < r_2 < \dots < r_K$ . If  $\rho_K = \frac{\lambda}{\mu_K} < 1$ , one may show that a stationary optimal policy exists by proving lemmas corresponding to Lemmas 4.1, 4.2, and 4.3 in essentially the same way. One may also include rewards for completed service.

#### 4.2. Qualitative Attributes of an Optimal Policy

We now investigate properties of  $V_f^*$ , an optimal cost function, and  $f^*$ , an optimal stationary policy. The cost function  $V_f^*$  must satisfy

$$\begin{aligned} V_f^*(i,1) = \min & \left[ \frac{\lambda}{\lambda+\mu_1} V_f^*(i+1,1) + \frac{\mu_1}{\lambda+\mu_1} V_f^*(i-1,1) - \frac{\phi_f^*}{\lambda+\mu_1} \right. \\ & \left. + \frac{c_i+r_1}{\lambda+\mu_1}, V_f^*(i,2) \right] \quad i = 1, 2, \dots \end{aligned}$$

$$v_f^*(0,1) = \min[v_f^*(1,1) - \frac{\phi_f^*}{\lambda} + \frac{r_1}{\lambda}, v_f^*(0,2)]$$

$$v_f^*(i,2) = \min[\frac{\lambda}{\lambda+\mu_2} v_f^*(i+1,2) + \frac{\mu_2}{\lambda+\mu_2} v_f^*(i-1,2)$$

$$- \frac{\phi_f^*}{\lambda+\mu_2} + \frac{c_i+r_2}{\lambda+\mu_2}, v_f^*(i,1)]$$

$$v_f^*(0,2) = \min[v_f^*(1,2) = \frac{\phi_f^*}{\lambda} + \frac{r_2}{\lambda}, v_f^*(0,1)].$$

Since  $v_f^*(i,1) \leq v_f^*(i,2) \leq v_f^*(i,1)$ ,  $v_f^*(i,1) = v_f^*(i,2)$  and the above functional equations may be rewritten as

$$v_f^*(i) = \min_{k=1,2} [\frac{\lambda}{\lambda+\mu_k} v_f^*(i+1) + \frac{\mu_k}{\lambda+\mu_k} v_f^*(i-1,1) - \frac{\phi_f^*}{\lambda+\mu_k} + \frac{c_i+r_k}{\lambda+\mu_k}]$$

$i = 1, 2, \dots$

$$v_f^*(0) = \min_{k=1,2} [v_f^*(1) - \frac{\phi_f^*}{\lambda} + \frac{r_k}{\lambda}].$$

We see that if the minimum is attained for  $k = 1$ , then  $f^*(i) = 1$ . Any  $i$  for which this is true will be called a slow service point. If the minimum is attained for  $k = 2$ , then  $f^*(i) = 2$ . Any  $i$  for which this is true will be called a fast service point. The determination of an optimal policy is equivalent to optimally classifying each  $i$  as slow or fast. We shall agree that if for some  $i$  the minimum is attained for both  $k = 1$  and  $k = 2$ , we shall set  $f^*(i) = 1$ .

We now prove a number of lemmas which correspond to those of Crabbill [3], Chapter II. The difference is that there is never any reference to truncated problems. The lemma of Crabbill's which most nearly corresponds

to each of the following lemmas will be noted. It should also be pointed out that the inequality manipulations are much the same as Crabbill's.

LEMMA 4.5. [Crabbill Lemma 1]:  $f^*(0) = 1$ .

Proof. Since  $r_1 < r_2$ ,

$$v_{f^*}(1) - \frac{\phi_{f^*}}{\lambda} + \frac{r_1}{\lambda} < v_{f^*}(1) - \frac{\phi_{f^*}}{\lambda} + \frac{r_2}{\lambda} .$$

LEMMA 4.6. [Crabbill Lemma 2]: Letting  $R = \frac{r_2 - r_1}{\mu_2 - \mu_1}$ ,

$$v_{f^*}(i) - v_{f^*}(i-1) > R \Leftrightarrow f^*(i) = 2 \quad i = 1, 2, \dots$$

$$v_{f^*}(i) - v_{f^*}(i-1) \leq R \Leftrightarrow f^*(i) = 1 \quad i = 1, 2, \dots$$

Proof. This follows immediately from the functional equation defining an optimal policy.

LEMMA 4.7. [Crabbill Lemma 3]: The set of  $i$  for which  $f^*(i) = 2$  is unbounded.

Proof. Assume the contrary; then there exists an  $N$  such that for all  $i \geq N$ ,  $f^*(i) = 1$  and

$$\mu_1(v_{f^*}(i) - v_{f^*}(i-1)) = \lambda(v_{f^*}(i+1) - v_{f^*}(i)) - \phi_{f^*} + c_i + r_1.$$

From Lemma 4.6

NWC TP 5594

$$\mu_1(v_{f^*}(i) - v_{f^*}(i-1)) \geq \lambda R - \phi_{f^*} + c_i + r_1.$$

Since  $\phi_{f^*} < \infty$  and  $c_i \rightarrow \infty$  as  $i \rightarrow \infty$ , one may find  $i_0 > N$  such that

$$\lambda R - \phi_{f^*} + c_{i_0} + r_1 > R\mu_1,$$

i.e.,

$$v_{f^*}(i_0) - v_{f^*}(i_0-1) > R.$$

This result contradicts Lemma 4.6.

LEMMA 4.8. [Crabill Lemmas 4 & 6]: If  $f^*(i) = 2$ , then

$$f^*(i+k) = 2 \text{ for } k = 1, 2, \dots .$$

Proof. The proof of this lemma is virtually identical to Crabill's Lemma 4.

Assume  $f^*(i) = 2$  and  $f^*(i+1) = 1$ .

$$f^*(i) = 2 \Rightarrow v_{f^*}(i) - v_{f^*}(i-1) > R$$

$$F^*(i+1) = 1 \Rightarrow v_{f^*}(i+1) - v_{f^*}(i) \leq R.$$

For  $i$  we have

$$\phi_{f^*} = c_i + r_2 - \mu_2(v_{f^*}(i) - v_{f^*}(i-1)) + \lambda(v_{f^*}(i+1) - v_{f^*}(i))$$

and

$$\phi_{f^*} < c_i + r_2 - (\mu_2 - \lambda)R.$$

For  $i+1$  we have

$$\phi_{f^*} = c_{i+1} + r_1 - \mu_1 (v_{f^*}(i+1) - v_{f^*}(i)) + \lambda (v_{f^*}(i+2) - v_{f^*}(i+1))$$

and

$$\phi_{f^*} \geq c_{i+1} + r_1 - \mu_1 R + \lambda (v_{f^*}(i+2) - v_{f^*}(i+1)).$$

It follows that

$$\lambda (v_{f^*}(i+2) - v_{f^*}(i+1)) < \lambda R - (c_{i+1} - c_i)$$

and

$$v_{f^*}(i+2) - v_{f^*}(i+1) < R$$

$$\Rightarrow f^*(i+2) = 1.$$

For  $i+2$  we have

$$\phi_{f^*} = c_{i+2} + r_1 - \mu_1 (v_{f^*}(i+2) - v_{f^*}(i+1)) + \lambda (v_{f^*}(i+3) - v_{f^*}(i+2))$$

and

$$\phi_{f^*} \geq c_{i+2} + r_1 - \mu_1 R + \lambda (v_{f^*}(i+3) - v_{f^*}(i+2)).$$

It follows that

$$\begin{aligned} \lambda (v_{f^*}(i+3) - v_{f^*}(i+2)) &\leq \phi_{f^*} - c_{i+2} - r_1 + \mu_1 R \\ &\leq \lambda R - (c_{i+2} - c_i). \end{aligned}$$

This implies

$$f^*(i+3) = 1.$$

Proceeding in this same way,  $f^*(i+k) = 1$  for  $K = 2, 3, 4, \dots$ . This result contradicts Lemma 4.7.

We are now in a position to state a stronger theorem than that of Crabill.

**THEOREM 4.9. [Crabill Theorem 1]**

If  $\sum_{i=0}^{\infty} \rho_2 c_i < \infty$ , then a stationary optimal policy  $f^*$  exists and is characterized by a single positive finite integer  $N$ , such that  $f^*(i) = 1$  for  $i < N$  and  $f^*(i) = 2$  for  $i \geq N$ .

Proof. The proof follows as a consequence of Theorem 4.4, Lemma 4.5, and Lemma 4.8.

The approach to the control of the M/M/1 queue with variable service rate presented here is formulated as a semi-Markov decision process, whereas Crabill formulated the problem as a continuous time Markov decision process. The semi-Markov formulation allows for instantaneous changes in state in a natural way, whereas instantaneous changes in state present some conceptual problems in the continuous time formulation.

Since this result shows that the optimal stationary policy is optimal over all admissible policies, the remarks following Theorem 4.4 show that Crabill's optimal stationary policy for the  $K$  service rate problem is an optimal policy. The ideas presented in this section may be combined

with those of Crabbill to extend the results to  $K$  service rates. In particular, proofs that do not depend on queue truncation may be made.

#### 4.3. Quantitative Results for the Linear Holding Cost Case

In the linear holding cost case we have  $c_i = h_i$  where  $h$  is the cost per unit time of holding a customer in the queueing system. Based on the preceding qualitative results, we are led to the following system of difference equations to obtain relative costs and  $\phi_N$ , the expected average cost per unit time associated with a policy of the form described in Theorem 4.9:

$$\lambda v_0 - \lambda v_1 = r_1 - \phi_N$$

$$(\lambda + \mu_1)v_i - \mu_1 v_{i-1} - \lambda v_{i+1} = h_i + r_1 - \phi_N, \quad 0 < i < N$$

$$(\lambda + \mu_2)v_i - \mu_2 v_{i-1} - \lambda v_{i+1} = h_i + r_2 - \phi_N \quad i \geq N$$

For  $i \geq N$  we set

$$v_i = A_2 + B_2 i + C_2 i^2.$$

For  $i < N$  we make use of the homogenous solution and have

$$v_i = B_1 i + C_1 i^2 + K \left[ \left( \frac{\mu_1}{\lambda} \right)^i - 1 \right],$$

which makes  $v_0 = 0$ . One easily finds that

$$B_i = \frac{h(\mu_i + \lambda)}{2(\mu_i - \lambda)^2} + \frac{r_i - \phi_N}{\mu_i - \lambda} \quad i = 1, 2$$

$$c_i = \frac{h}{2(\mu_i - \lambda)} \quad i = 1, 2.$$

The above system must satisfy the boundary conditions

$$A_2 + B_2 N + C_2 N^2 = B_1 N + C_1 N^2 + K \left[ \left( \frac{\mu_1}{\lambda} \right)^N - 1 \right]$$

$$A_2 + B_2(N-1) + C_2(N-1)^2 = B_1(N-1)^2 + C_1(N-1)^2 + K \left[ \left( \frac{\mu_1}{\lambda} \right)^{N-1} - 1 \right].$$

Finally,

$$\begin{aligned} \phi_N &= \lambda v_1 + r_1 \\ &= \lambda (B_1 + C_1) + \lambda K \left[ \left( \frac{\mu_1}{\lambda} \right)^N - 1 \right] + r_1. \end{aligned}$$

Elimination of  $A_2$  in the boundary condition equation and substitution for  $B_1$ ,  $C_1$ ,  $B_2$ , and  $C_2$  gives

$$\phi_N - \frac{K(\mu_2 - \lambda)(\mu_1 - \lambda)^2}{\lambda(\mu_2 - \mu_1)} \left( \frac{\mu_1}{\lambda} \right)^{N-1} = \frac{\rho_2 h}{1-\rho_2} + r_2 + \frac{h}{1-\rho_1} + h(N-1) - \frac{(r_2 - r_1)(\mu_2 - \lambda)}{(\mu_2 - \mu_1)}.$$

Substitution for  $B_1 + C_1$  in the expression for  $\phi_N$  yields

$$\phi_N - \frac{K(\mu_1 - \lambda)^2}{\mu_1} = \frac{\rho_1 h}{1-\rho_1} + r_1$$

where  $\rho_i = \frac{\lambda}{\mu_i}$ . Elimination of  $K$  gives

$$\phi_N = \frac{\rho_1 h}{1-\rho_1} + r_1 + \frac{\frac{\rho_2 h}{1-\rho_2} + r_2 + hN + \frac{r_1(\mu_1 - \lambda) - r_2(\mu_2 - \lambda)}{(\mu_2 - \mu_1)}}{1 - \frac{(\mu_2 - \lambda)(\mu_1)}{(\mu_2 - \mu_1)(\lambda)}^N},$$

One may verify that

$$\phi_0 = \frac{\rho_2 h}{1-\rho_2} + r_2 \quad (\text{Always use } \mu_2)$$

$$\phi_1 = \frac{\rho_2 h}{1-\rho_2} + \rho_2 r_2 + (1-\rho_2)r_1 \quad (\text{Use } \mu_1 \text{ only for } i = 0)$$

and providing  $\lambda < \mu_1$ ,

$$\phi_\infty = \frac{\rho_1 h}{1-\rho_1} + r_1 \quad (\text{Always use } \mu_1).$$

Setting

$$G(N) = \frac{\phi_0 + hN + \frac{r_1(\mu_1 - \lambda) - r_2(\mu_2 - \lambda)}{\mu_2 - \mu_1}}{\frac{(\mu_2 - \lambda)^N}{(\mu_2 - \mu_1)} \left(\frac{\mu_1}{\lambda}\right) - 1},$$

then

$$\phi(n) = \phi_\infty - G(N).$$

It is conjectured in Crabil [3] that  $\phi$  is a convex function of  $N$ .

We now show that  $\phi$  is not convex in  $N$ . A necessary and sufficient condition for  $\phi$  to be convex in  $N$  is that

$$\frac{d^2 G(N)}{dN^2} \leq 0 \quad \text{for all } N.$$

Consider the case  $\lambda < \mu_1$ . For convenience we write

$$G(N) = \frac{A + hN}{k_1 \beta^N - 1}$$

where

$$k_1 = (\mu_2 - \lambda) / (\mu_2 - \mu_1) > 1$$

$$\beta = \frac{\mu_1}{\lambda} > 1$$

$$A = \phi_0 + \frac{r_1(\mu_1 - \lambda) - r_2(\mu_2 - \lambda)}{\mu_2 - \mu_1},$$

then

$$\begin{aligned} G'(N) &= \frac{(k_1 \beta^N - 1)h - (A + hN)k_1 \beta^N \log \beta}{(k_1 \beta^N - 1)^2} \\ &= \frac{h - G(N)k_1 \beta^N \log \beta}{(k_1 \beta^N - 1)}, \end{aligned}$$

and

$$\begin{aligned} G''(N) &= -(k_1 \beta^N - 1) \left[ \frac{G(N)k_1 \beta^N (\log \beta)^2 + k_1 \beta^N \log \beta G'(N)}{(k_1 \beta^N - 1)^2} \right. \\ &\quad \left. - \frac{(h - G(N)k_1 \beta^N \log \beta)k_1 \beta^N \log \beta}{(k_1 \beta^N - 1)^2} \right] \end{aligned}$$

which after some manipulation reduces to

$$G''(N) = \frac{k_1 \beta^N \log \beta}{(k_1 \beta^N - 1)^2} \left[ -2h + (A + hN) \log \frac{\mu_1}{\lambda} \frac{(1 + k_1 \beta^N)}{(-1 + k_1 \beta^N)} \right].$$

Since  $A + hN > 0$  for  $N$  sufficiently large

$$G''(N) > \frac{k_1 \beta^N \log \beta}{(k_1 \beta^N - 1)^2} [-2h + \log \frac{\mu_1}{\lambda} (A + hN)]$$

for  $N$  sufficiently large. Also for  $N$  sufficiently large the expression in brackets is positive, so  $\phi(N)$  is not convex.

However,  $\phi(N)$  is unimodal, since  $G(N)$  is unimodal. To show this when  $\lambda < \mu_1$ , consider  $G(N)$ . Since  $A + hN > 0$  for  $N$  sufficiently large,  $G(N) > 0$  for  $N$  sufficiently large.

Moreover,

$$\lim_{N \rightarrow \infty} G(N) = 0,$$

showing that  $G$  is decreasing for sufficiently large  $N$ . Since

$$G(1) - G(0) = \phi(0) - \phi(1) > 0,$$

$G$  is increasing for some positive  $N$  and hence  $G$  has at least one relative maximum for positive values of  $N$ .

A necessary condition for a relative maximum is that  $G'(N) = 0$ .

Thus we are interested in  $N$  for which

$$A + hN = \frac{(k_1 \beta^N - 1)h}{k_1 \beta^N \log \beta}.$$

We observe that the expression on the right as a function of  $N$  is non-negative, increasing and strictly concave. At  $N = 0$  this function has value  $(1 - 1/k_1) \frac{h}{\log \beta}$  and as  $N \rightarrow \infty$  the function approaches  $h/\log \beta$ . The derivative at  $N = 0$  is  $h/k_1$ . The expression on the left is linear in  $N$  with value  $A$  at  $N = 0$  and derivative  $h$  at  $N = 0$ . Since  $G$

has at least one relative maximum for positive  $N$ ,

$$(1 - 1/k_1) \frac{h}{\log \beta} > A.$$

If this were not the case, there would be no positive  $N$  satisfying the condition  $G'(n) = 0$ . This follows from the concavity of the expression on the right and the fact that slopes at zero are such that

$$h/k_1 < h.$$

We may plot both expressions for  $N$  to graphically solve for minimizing values of  $N$ . A typical plot is given in Figure 1.

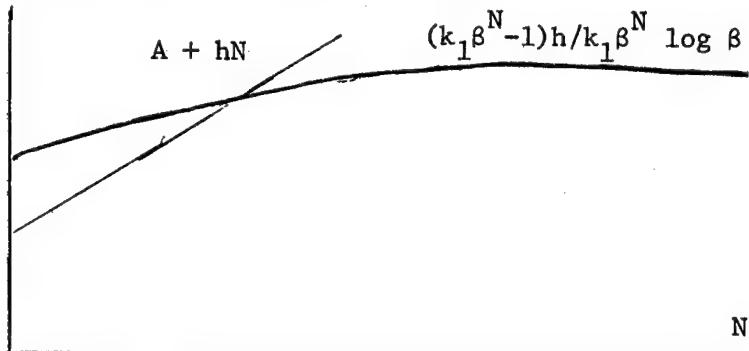


Figure 1. Graphical solution for  $G'(N) = 0$ .

From the concavity of the expression on the right and the linearity of the expression on the left the  $N$  for which  $G'(N) = 0$  is unique, so  $G(N)$  has exactly one relative maximum for positive  $N$ .

The value of  $N$  which maximizes  $G$  may be obtained by successive approximation. We begin by solving

$$A + hx_0 = (1 - 1/k_1) \frac{h}{\log \beta}$$

$$\vdots$$

$$A + hx_i = \frac{(k_1 \beta^{x_{i-1}} - 1)h}{k_1 \beta^{x_{i-1}} \log \beta} .$$

A graphical interpretation of this procedure is given in Figure 2.

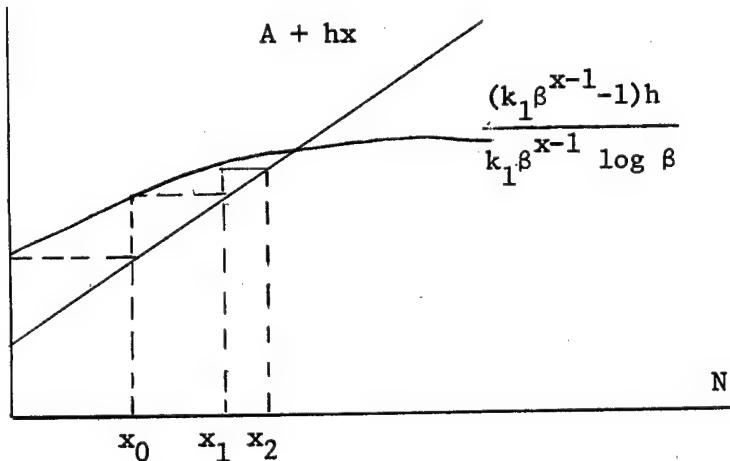


Figure 2. Successive approximation solution for optimal  $N$ .

If  $\mu_1 \leq \lambda$ , unimodality may be established similarly but the method of successive approximations will diverge, so one must solve the transcendental equation associated with a relative minimum differently or use a search technique to minimize  $\phi_N$  directly. Since  $\phi_N$  is unimodal, standard procedures may be used. Another possibility is to use the policy improvement algorithm. With this algorithm, if one begins with a policy such that  $f(i) = 2$  for  $i \geq N$  and  $f(i) = 1$  for  $i < N$ , then all improvements will be of this form and the method converges to the optimal value of  $N$ .

## Appendix A

## FUNCTIONAL EQUATIONS OF OPTIMALITY

Introduction

This appendix is written to display the functional equations of optimality for special cases that arise in the queueing applications in this report. We recall that a stationary semi-Markov process starts at time 0 in state  $i = 0, 1, \dots$  with probability  $P_i$ . With the observation of state  $i$  an action  $k = 1, 2, \dots$ , or  $K$  is taken. The next state,  $j$ , of the process occurs with probability  $P_{ij}(k)$ . Conditional on the event that the next state is  $j$ , the time until the transition from  $i$  to  $j$  occurs is a random variable with distribution function  $F_{ij}(\cdot | k)$ . With the observation of state  $j$  an action  $k = 1, 2, \dots K$  is taken and this procedure goes on indefinitely. Whenever the process is in state  $i$  and action  $k$  is taken, a cost is incurred which depends on random events occurring during the transition interval.

We are interested in two optimization criteria, minimum expected average cost per unit time and minimum expected total discounted cost. For the average cost criterion the functional equation of optimality is given by

$$V_{f^*}(i) = \min_k \{C_i(k) - \phi_{f^*} t_i(k) + \sum_j P_{ij}(k) V_{f^*}(j)\}$$

where  $C_i(k)$  is the expected cost of a transition and  $t_i(k)$  is the expected transition time when action  $k$  is taken in state  $i$ . The

expected average cost per unit time,  $\phi_{f*}$ , is obtained from  $V_{f*}(0) = 0$ .

The functional equation of optimality for the discounted cost criterion is given by

$$V_\beta(i) = \min_k \{C_i(k) + \sum_{j=0}^{\infty} P_{ij}(k) \int_0^{\infty} e^{-\beta t} V_\beta(j) dF_{ij}(t|k)\}$$

where  $C_i(k)$  is the expected discounted cost of a transition when action  $k$  is taken in state  $i$ , and the discount factor  $\beta$  implies a cost  $C$  at time  $t$  contributes  $C e^{-\beta t}$  to the total discounted cost.

#### Functional Equations in a Special Case

In this section we consider the form of the above functional equations for a special case that arises in the applications. First, we assume that

$$F_i(t|k) = \sum_j P_{ij}(k) F_{ij}(t|k)$$

is basic in the problem formulation. We also assume that if action  $k$  is taken in state  $i$  and the transition time is  $t$  then  $j$ , the next state of the process, is given by

$$j = i + m + s$$

where  $m$  has a Poisson distribution with parameter  $\lambda_i(k)t$ , and  $s$  is independent of  $m$  and  $t$  with finite distribution  $P_i(s|k)$ . It is understood that  $s$  may take on negative as well as positive integer values. We set

$$p(n;\lambda) = \frac{\lambda^n}{n!} e^{-\lambda}$$

and our problem is to express  $P_{ij}(k)$  and  $F_{ij}(t|k)$  in terms of  $F_i(t|k)$ ,  $P_i(s|k)$  and  $p(n;\lambda_i(k)t)$ .

For fixed  $t$

$$P_{ij}(k|t) = \sum_v P_i(s+v|k)p(m-v; \lambda_i(k)t)$$

and

$$P_{ij}(k) = \sum_v P_i(s+v|k) \int_0^\infty p(m-v; \lambda_i(k)t) dF_j(t|k)$$

The joint event, the next state is  $j$  and the transition time is less than  $t$  given state  $i$  and decision  $k$ , has probability

$$P[j \text{ and } T \leq t] = \sum_v P_i(s+v|k) \int_0^t p(m-v; \lambda_i(k)x) dF_j(x|k)$$

It follows that

$$F_{ij}(t|k) = \frac{\sum_v P_i(s+v|k) \int_0^t p(m-v; \lambda_i(k)x) dF_j(x|k)}{\sum_v P_i(s+v|k) \int_0^\infty p(m-v; \lambda_i(k)x) dF_j(x|k)}$$

For the average cost criterion the functional equation of optimality becomes

$$\begin{aligned} v_{f*}(i) &= \min_k \{C_i(k) - \phi_{f*} t_i(k) \\ &\quad + \sum_{s,m,v} P_i(s+v|k) v_{f*}(i+m+s) \int_0^\infty p(m-v; \lambda_i(k)t) dF_j(t|k)\} \end{aligned}$$

For the discounted cost criterion we have

$$\begin{aligned} v_\beta(i) &= \min_k \{C_i(k) \\ &\quad + \sum_{s,m,v} P_i(s+v|k) v_\beta(i+m+s) \int_0^\infty e^{-\beta t} p(m-v; \lambda_i(k)t) dF_j(t|k)\} \end{aligned}$$

If a particular action  $k$  is taken in state  $i$  then  $F_i(t|k)$ ,  $P_i(s|k)$ , and  $\lambda_i(k)$  will be determined for a given  $i$  and  $k$ . This consideration allows one to write in more detail the entries for particular  $i$  and  $k$  on the right hand sides of the functional equations above. There are essentially two classes of actions used in this report.

If action  $k$  belongs to class 1 and it is taken in state  $i$  then the transition time is exponential and the transition rate from state  $i$  to state  $j$  is  $\alpha_{ij}(k)$ . In this case

$$F_i(t|k) = 1 - e^{-\alpha_i(k)t}$$

where

$$\alpha_i(k) = \sum_j \alpha_{ij}(k)$$

For actions of class 1 it is further assumed that  $\lambda_i(k) = 0$ . It follows that in the equation  $j = i + m + s$ ,  $m = 0$  with probability 1, and

$$P_i(j-i|k) = \frac{\alpha_{ij}(k)}{\alpha_i(k)} = P_{ij}(k)$$

For the average cost criterion in this case

$$t_i(k) = \int t dF_i(t|k) = \frac{1}{\alpha_i(k)}$$

and

$$\sum_s \sum_m \sum_v P_i(s+v|k) V_{f^*}(i+m+s) \int_0^\infty p(m-v; \lambda_i(k)t) dF_i(t|k)$$

$$= \sum_j \frac{\alpha_{ij}(k)}{\alpha_i(k)} V_{f^*}(j)$$

For the discounted cost criterion in this case,

$$\begin{aligned}
 & \sum_{s \in V} \sum_{m \in V} P_i(s+v|k) v_\beta(i+m+s) \int_0^\infty e^{-\beta t} p(m-v; \lambda_i(k)t) dF_i(t|k) \\
 &= \sum_j \frac{\alpha_{ij}(k)}{\alpha_i(k)} v_\beta(j) \int_0^\infty e^{-\beta t} \alpha_i(k) e^{-\alpha_i(k)t} dt \\
 &= \sum_j \frac{\alpha_{ij}(k)}{\alpha_i(k) + \beta} v_\beta(j)
 \end{aligned}$$

If action  $k$  belongs to class 2 and it is taken in state  $i$  then  $F_i(t|k)$  is general,  $\lambda_i(k) > 0$ , and there exists an  $s_0$  such that  $P_i(s_0|k) = 1$ . For example,  $s_0 = -1$  and  $s_0 = -i$  are important special cases. We have for the average cost criterion in this case

$$t_i(k) = \int t dF_i(t|k) \equiv \mu_i^{-1}(k)$$

and

$$\begin{aligned}
 & \sum_{s \in V} \sum_{m \in V} P_i(s+v|k) v_{f*}(i+m+s) \int_0^\infty p(m-v; \lambda_i(k)t) dF_i(t|k) \\
 &= \sum_m v_{f*}(i+m+s_0) \int_0^\infty p(m; \lambda_i(k)t) dF_i(t|k)
 \end{aligned}$$

For the discounted cost criterion in this case,

$$\begin{aligned}
 & \sum_{s \in V} \sum_{m \in V} P_i(s+v|k) v_\beta(i+m+s) \int_0^\infty e^{-\beta t} p(m-v; \lambda_i(k)t) dF_i(t|k) \\
 &= \sum_m v_\beta(i+m+s_0) \int_0^\infty e^{-\beta t} p(m; \lambda_i(k)t) dF_i(t|k)
 \end{aligned}$$

If, moreover,

$$F_1(t|k) = 0 \quad t < 0$$

$$F_1(t|k) = 1 \quad t \geq 0$$

then the right hand side above is simply  $V_{f^*}(i+s_0)$  for the average cost criterion and  $V_\beta(i+s_0)$  for the discounted cost criterion.

#### Expected Transition Costs

To complete the specification of the functional equations of optimality we must compute  $C_i(k)$ , the expected cost of a transition if action  $k$  is taken in state  $i$ . For the discounted cost criterion this calculation is required for all  $i$  and  $k$  assuming that action  $k$  is taken at time 0. Since the calculation for the average cost criterion is the same regardless of the time of the action it is convenient to always think of action  $k$  taken in state  $i$  at time 0.

We now derive  $C_i(k)$  when the cost structure is linear. If at time 0 the process is in  $i$  and action  $k$  is taken, an instantaneous cost  $c_{ik}$  is incurred and costs start to accrue at a rate of  $r_k + h_k i$ . Costs accrue at this rate until either the transition ends or the occurrence of a random event at time  $T_1$ . If  $T_1$  occurs before the end of the transition then beginning at time  $T_1$  costs accrue at a rate of  $r_k + h_k(i+1)$ . Thus if  $m$  random events occur at times  $0 < T_1 < T_2 \dots < T_m < t$  during the transition time interval  $(0, t)$ , the cost rate over time interval  $(T_n, T_{n+1})$  is  $r_k + h_k(i+n)$  for  $n = 0, 1, \dots, m$  where  $T_0 = 0$  and  $T_{m+1} = t$ . We assume that  $T_1, T_2, \dots$  are generated according to a Poisson process with parameter  $\lambda_i(k)$ .

It follows that for a transition time  $t$  the number of random events occurring at times  $T_1 < T_2 < \dots < t$  has a Poisson distribution with parameter  $\lambda_i(k)t$ .

We recall that if action  $k$  is taken in state  $i$  and the transition time is  $t$  then  $j$ , the next state of the process, satisfies

$$j = i + m + s$$

where  $m$  has a Poisson distribution with parameter  $\lambda_i(k)t$ . Whereas only the number of these events is important in determining the next state of the process, both the number and times of occurrence of these events determine transition costs.

For the average cost criterion the expected transition cost if action  $k$  is taken in state  $i$ , given  $T_1, T_2, \dots, T_m, m$ , and  $t$  is

$$\begin{aligned} c_{ik}(k|T_1, T_2, \dots, T_m, m, t) \\ = c_{ik} + \sum_{n=0}^m (T_{n+1} - T_n)(r_k + h_k(i+n)) \\ = c_{ik} + (r_k + h_k i)t + h_k m t - h_k \sum_{i=1}^m T_i \end{aligned}$$

Similarly for the discounted cost criterion

$$\begin{aligned} c_{ik}(k|T_1, T_2, \dots, T_m, m, t) \\ = c_{ik} + \sum_{n=0}^m (r_k + h_k(i+n)) \int_{T_n}^{T_{n+1}} e^{-\beta s} ds \\ = c_{ik} + (r_k + h_k i) \int_0^t e^{-\beta s} ds + \sum_{n=0}^m h_k n \int_{T_n}^{T_{n+1}} e^{-\beta s} ds \\ = c_{ik} + \left( (r_k + h_k i)(1 - e^{-\beta t}) - mh_k e^{-\beta t} + h_k \sum_{i=1}^m e^{-\beta T_i} \right) / \beta \end{aligned}$$

Conditional on  $m$  and  $t$ ,  $T_1, T_2, \dots, T_m$  are independently and uniformly distributed over  $(0, t)$ . It follows that

$$E \sum_{i=1}^m T_i = \frac{m}{t} \int_0^t x dx = \frac{mt}{2}$$

and

$$E \sum_{i=1}^m e^{-\beta T_i} = \frac{m}{t} \int_0^t e^{-\beta x} dx = \frac{m}{\beta t} (1 - e^{-\beta t})$$

The conditional cost of action  $k$  in state  $i$  given  $m$  and  $t$  is

$$c_i(k|m, t) = c_{ik} + (r_k + h_k i)t + h_k \frac{mt}{2}$$

for the average cost criterion and

$$c_i(k|m, t) = c_{ik} + \frac{(h_k i + r_k)}{\beta} (1 - e^{-\beta t})$$

$$+ \frac{mh_k}{\beta^2 t} (1 - e^{-\beta t} - \beta t e^{-\beta t})$$

for the discounted cost criterion.

Since

$$E(m|t) = \lambda_i(k)t,$$

we have the expected transition cost for the average cost criterion

$$c_i(k) = c_{ik} + (r_k + h_k i)\mu_i^{-1}(k) + h_k \frac{\lambda_i(k)}{2} \mu_{i2}(k)$$

where

$$\mu_i^{-1}(k) = \int t dF_i(t|k)$$

and

$$\mu_{i2}(k) = \int t^2 dF_i(t|k).$$

For the discounted cost criterion

$$C_i(k) = C_{ik} + \frac{(h_k i + r_k)}{\beta} (1 - \int_0^{\infty} e^{-\beta t} dF_i(t|k)) \\ + \frac{\lambda_i(k) h_k}{\beta^2} (1 - \int_0^{\infty} e^{-\beta t} dF_i(t|k) - \beta \int_0^{\infty} t e^{-\beta t} dF_i(t|k))$$

For actions in class 1 where  $\lambda_i(k) = 0$  and  $F_i(t|k) = 1 - e^{-\alpha_i(k)t}$

the expected transition cost for the average cost criterion is

$$C_i(k) = C_{ik} + \frac{(r_k + h_k i)}{\alpha_i(k)}$$

For the discounted cost criterion in this case

$$C_i(k) = C_{ik} + \frac{(r_k + h_k i)}{\alpha_i(k) + \beta}$$

For actions in class 2 when  $\lambda_i(k) > 0$  there is no special reduction of the above formulae. If, however,

$$F_i(t|k) = 0 \quad t < 0$$

$$= 1 \quad t \geq 0$$

$$C_i(k) = C_{ik}$$

for both the average cost and discounted cost criterion.

## Appendix B

## GLOSSARY

This glossary summarizes the notation and key results of Reed [7].

Section A provides basic concepts. Section B is concerned with the case when the optimization criterion is minimum expected discounted costs.

Section C is concerned with the case when the optimization criterion is minimum expected average cost per unit time. All theorem numbers refer to theorems in Reed [7].

A. Basic Concepts and Notation

Semi-Markov decision process: A sequential decision process associated with a semi-Markov process which starts at time 0 in one of the states  $i = 0, 1, 2, \dots$ . With the observation of state  $i$ , an action  $k = 1, 2, \dots$ , or  $K$  is taken. The next state,  $j$ , of the process occurs with probability  $P_{ij}(k)$ . Conditional on the event that the next state is  $j$ , the time until the transition from  $i$  to  $j$  occurs is a random variable with distribution function  $F_{ij}(\cdot|k)$ . With the observation of state  $j$ , an action  $k = 1, 2, \dots$ , or  $K$  is taken, and this iterative procedure goes on indefinitely. The transition time distribution when action  $k$  is taken in state  $i$  is given by

$$F_i(t|k) = \sum_j P_{ij}(k) F_{ij}(t|k).$$

Trivial sequence of decisions: A sequence of decisions  $k_1, k_2, \dots, k_m$  is said to be trivial with respect to state  $i$  if

$$P_{i_r i_{r+1}}(k_r) = 1 \quad r = 1, 2, \dots, m$$

and

$$\begin{aligned} F_{i_r i_{r+1}}(t|k_r) &= 0 & t < 0 \\ &= 1 & t \geq 0 \end{aligned}$$

for  $r = 1, 2, \dots, m$ , where  $i_1 = i_{m+1} = i$  and  $i_r \neq i$  for  $r = 2, 3, \dots, m$ .

Condition 1: There exists an integer  $n_0$  and a fraction  $q$  with  $0 < q \leq 1$  such that for all  $n \geq n_0$  with decisions  $k_1, k_2, \dots, k_n$  in states  $i_1, i_2, \dots, i_n$  and transition time distribution  $F_{i_1 i_2}(t|k_1), \dots, F_{i_n i_{n+1}}(t|k_n)$  there exists an  $\varepsilon > 0$  and a  $\delta > 0$  such that for  $j_1, j_2, \dots, j_n^*$

$$F_{i_{j_1} i_{j_1+1}}(\delta|k_{j_1}) \leq 1 - \varepsilon$$

where  $j_1, j_2, \dots, j_n^*$  is a subsequence of  $1, 2, \dots, n$  and  $n^* \geq qn$ .

M/M/1 Queue: A single server queue where customers arrive according to a Poisson process with parameter  $\lambda$ . Service times are independently and identically distributed with cumulative distribution function  $1 - e^{-\mu t}$ .

M/G/1 Queue: A single server queue where customers arrive according to a Poisson process with parameter  $\lambda$ . Service times are independently and identically distributed with cumulative distribution function  $B$ .

List of Symbols:

$\mu^{-1}$	Mean of service time distribution
$\mu_2$	Second moment about the origin of service time distribution
$\rho = \lambda \mu^{-1}$	Queue utilization factor
$\Pi$	Class of admissible policies
$\pi$	Element of $\Pi$
$\mathcal{F}$	Class of Stationary Policies
$f$	Element of $\mathcal{F}$

B. List of Symbols and Key Theorems for Discounted Cost CriterionList of Symbols:

$\beta$	Discount factor
$C e^{-\beta t}$	Equivalent cost at time 0 of a cost $C$ incurred at time $t$
$C_i(k)$	Expected discounted cost of a transition when action $k$ is taken in state $i$
$V_\pi(i)$	Total expected discounted cost of using $\pi$ given the process begins at time 0 in state $i$
$V_\beta(i)$	Optimal discounted cost function where
	$V_\beta(i) = \inf_{\pi \in \Pi} V_\pi(i)$
$\pi^*$	$\pi^*$ is $\beta$ -optimal if $V_{\pi^*}(i) = V_\beta(i)$ for all $i$
$f^*$	$f^*$ is stationary $\beta$ -optimal if $f^* \in \mathcal{F}$ and $V_{f^*}(i) = V_\beta(i)$ for all $i$

Key Theorems:

THEOREM 3.10. (Functional Equation of Optimality)

$$V_\beta(i) = \min_k \{ C_i(k) + \sum_{j=0}^{\infty} P_{ij}(k) \int_0^{\infty} e^{-\beta t} V_\beta(j) dF_{ij}(t|k) \} \quad (3.5)$$

(If  $v_f(i)$  satisfies (3.5), then  $f$  is called a  $\beta$ -optimal improvement policy.)

## THEOREM 3.12.

If  $c_i(k) \geq 0$  for all  $i$  and  $k$ , then there exists a stationary  $\beta$ -optimal policy.

## THEOREM 3.17.

If the condition of Theorem 3.12 holds,  $v_{f^*}(i)$  satisfies (3.5) and  $v_{f^*}(i) \leq Q(i)$  where  $Q$  is a polynomial of finite degree  $r$ , increasing in  $i$  with

$$p_{i,h(i,k)+x}(k) = \int e^{-\lambda_k t} \frac{(\lambda_k t)^x}{x!} dF(t|k) \quad x = 0, 1, \dots; \\ k = 1, 2, \dots, K_0$$

where

$$\int t^r dF(t|k) < \infty \quad k = 1, 2, \dots, K_0$$

and

$$p_{i,h(i,k)+x}(k) = p_k(x) \quad x = 0, 1, \dots, M_1 \\ = 0 \quad x > M_1;$$

independent of  $F(t|k)$  for  $k = K_0 + 1, \dots, K$ , where  $h$  is a deterministic function of  $i$  and  $k$  satisfying

$$0 \leq h(i,k) \leq i + M_0 \quad k = 1, 2, \dots, K$$

then  $f^*$  is a  $\beta$ -optimal policy.

C. List of Symbols and Key Theorems for the Average Cost CaseList of Symbols:

$c_i(k)$	Expected cost of transition if action $k$ is taken in state $i$
$t_i(k)$	Expected time of transition if action $k$ is taken in state $i$
$\phi_{\pi}$	Expected average cost per unit time if policy $\pi$ is used
$\phi$	$\inf_{\pi \in \Pi} \phi_{\pi}$
$\pi^*$	$\pi^*$ is optimal if $\phi_{\pi^*} = \phi$
$f^*$	$f^*$ is stationary optimal if $f^* \in \mathcal{F}$ and $\phi_{f^*} = \phi$
$T_{i0}$	First passage time into state 0 given the process begins in state $i$
$C_{i0}$	Cost associated with first passage to state 0 given the process begins in state $i$
$V_{\pi}(i, \theta)$	Expected relative cost with respect to $\theta$ of first passage to state 0 given the process begins in state $i$ and policy $\pi$ is used, i.e.,
	$V_{\pi}(i, \theta) = E_{\pi}[C_{i0} - \theta T_{i0}]$
$V(i, \theta)$	Optimal relative cost function where

$$V(i, \theta) = \inf_{\pi \in \Pi} V_{\pi}(i, \theta)$$

Key Theorems:

THEOREM 4.1. (Functional Equation of Optimality)

For  $i = 0, 1, 2, \dots$ ,

$$V(i, \phi) = \min_k \{c_i(k) - \phi t_i(k) + \sum_{j=0}^{\infty} p_{ij}(k) V(j, \phi)\} \quad (4.3)$$

or in vector notation

$$V(\phi) = \min_{f \in \mathcal{F}} \{C(f) - \phi t(f) + P(f)V(\phi)\}$$

where  $\phi$  is determined from  $V(0, \phi) = 0$ . (If  $V_f(\phi_f)$  satisfies (4.3) with  $V_f(0, \phi_f) = 0$ , then  $f$  is called a 0-optimal improvement policy.)

## THEOREM 4.3.

If the  $C_i(k)$  are bounded below,  $\phi_\pi \leq M < \infty$  for some  $\pi$ , and  $S = \{i : C_i(k) - Mt_i(k) < 0 \text{ for some } k\}$  is finite, then  $V(\phi)$  exists and is bounded below.

Assumptions common to remaining theorems:

A(1):  $\Pi = \{\pi : \phi_\pi \leq M\}$  is non-empty where  $0 \leq M < \infty$ .

A(2): There exists a state, say state 0, that is positive recurrent over all  $\pi \in \Pi$ .

A(3):  $V(\phi)$  exists and is bounded below.

## THEOREM 4.4.

If A(1), A(2), and A(3) hold, then a stationary optimal policy  $f^*$  exists, and it satisfies the relationship

$$V_{f^*} = \min_f [C(f) - \phi_{f^*} t(f) + P(f)V_{f^*}] \quad (4.4)$$

where  $\phi_{f^*}$  is determined from  $V_{f^*}(0, \phi_{f^*}) = 0$ .

## THEOREM 4.8.

If assumptions A(1), A(2), A(3) hold and  $f^* \in \mathcal{F}$  is such that  $V_{f^*}$  and  $\phi_{f^*}$  satisfy (4.4) with  $x(f)V_{f^*} < \infty$  for all  $f \in \mathcal{F}$ , where  $x(f)$

is the stationary distribution of the imbedded Markov chain associated with  $f$ , then  $f^*$  is optimal over  $\Pi$ .

COROLLARY 4.9.

If assumptions A(1), A(2), A(3) hold and  $f^* \in \mathcal{F}$  is such that  $V_{f^*}$  and  $\phi_{f^*}$  satisfy (4.4) with  $V_{f^*} \leq M^*$ , then  $f^*$  is optimal over  $\Pi$ .

REFERENCES

1. Bell, C. E. (1971), "Characterization and Computation of Optimal Policies for Operating an M/G/1 Queueing System with Removable Server," Operations Research, 19, 208-218.
2. Blackburn, J. D. (1971), "Optimal Control of Queueing Systems with Intermittent Service," Technical Report No. 23, Department of Operations Research, Stanford University.
3. Crabil, T. B. (1969), "Optimal Control of a Queue with Variable Service Rates," Technical Report No. 75, Department of Operations Research, Cornell University. (A shorter version appears in Management Science, 18, 560-566.)
4. Heyman, D. P. (1968), "Optimal Operating Policies for M/G/1 Queueing Systems," Operations Research, 16, 362-382.
5. Howard, R. A. (1960), Dynamic Programming and Markov Processes, MIT Press, Cambridge, Mass.
6. Kakalik, J. S. (1969), "Optimal Dynamic Operating Policies for a Service Facility," Technical Report No. 47, Operations Research Center, Massachusetts Institute of Technology.
7. Reed, F. C. (1974), "Denumerable State Markov Decision Processes with Unbounded Costs," NWC TP 5593, Naval Weapons Center, China Lake, CA.

8. Ross, S. M. (1970), Applied Probability Models With Optimization Applications, Holden-Day, San Francisco.
9. Sabeti, H. (1970), "Optimal Decision in Queueing," ORC 70-12, Operations Research Center, University of California, Berkeley.
10. Sobel, M. J. (1969), "Optimal Average Cost Policy for a Queue with Start-up and Shut-down Costs," Operations Research, 17, 145-162.
11. Torbett, A. E. (1973), "Models for the Optimal Control of Markovian Closed Queueing Systems with Adjustable Service Rates," Technical Report No. 39, Department of Operations Research, Stanford University.

INITIAL DISTRIBUTION

1 Director of Navy Laboratories (MAT-03L)

10 Naval Air Systems Command

AIR-03B (1)

AIR-03P (1)

AIR-03P2 (1)

AIR-03P24, B. Glatt (1)

AIR-302 (2)

AIR-50174 (2)

AIR-5033 (1)

ADPO-31 (1)

3 Chief of Naval Operations

OP-098T (1)

OP-96 (1)

OP-96C (1)

4 Chief of Naval Material

MAT-03 (1)

MAT-03PB (1)

MAT-031 (1)

MAT-032 (1)

5 Naval Ordnance Systems Command

ORD-03 (1)

ORD-032 (1)

ORD-0632 (2)

PMO-403C (1)

7 Chief of Naval Research, Arlington

ONR-100 (1)

ONR-430 (2)

ONR-434

B. J. McDonald (1)

T. C. Varley (1)

ONR-460 (1)

ONR-462 (1)

1 Naval Air Development Center, Johnsville (Technical Library)

1 Naval Air Test Center, Patuxent River (Aeronautical Publications  
Library)

1 Naval Avionics Facility, Indianapolis (Technical Library)

1 Naval Electronics Laboratory Center, San Diego

1 Naval Observatory (Scientific Director)

2 Naval Ordnance Laboratory, White Oak

Mathematics Department (1)

Technical Library (1)

1 Naval Ordnance Station, Indian Head

6 Naval Postgraduate School, Monterey

- Code 02, Academic Dean (1)
- Code 0212, Librarian (1)
- Code 023, Dean of Research Administration (1)
- Code 53, Chairman, Department of Mathematics (2)
- Code 55, Chairman, Department of Operations Analysis (1)

2 Naval Research Laboratory

- Code 2620 (1)
- Code 7800 (1)

1 Naval Ship Research and Development Center, Annapolis (Bessie B. Day)

1 Naval Ship Research and Development Center, Bethesda

3 Naval Undersea Center, San Diego

- Code 14 (1)
- Code 133 (1)
- Code 3530, C. Sweat (1)

1 Naval Weapons Laboratory, Dahlgren (Technical Library)

1 Office of Naval Research Branch Office, Pasadena (R. Lau)

1 Operational Test and Evaluation Force

1 Pacific Missile Range (Technical Library)

1 Army Ballistics Research Laboratories, Aberdeen Proving Ground

1 Army Research Office, Durham

1 Harry Diamond Laboratories (Technical Library)

1 Aeronautical Systems Division, Wright-Patterson Air Force Base  
(Aeronautical Research Laboratory, Applied Mathematics Branch)

1 Air Force Office of Scientific Research, Arlington

1 Air Force Weapons Laboratory, Kirtland Air Force Base

1 Air University Library, Maxwell Air Force Base

2 Armed Forces Staff College, Norfolk

1 Defense Advanced Research Projects Agency, Arlington (Fred A. Koether)

12 Defense Documentation Center

1 National Aeronautics & Space Administration (RP)

1 Langley Research Center (Technical Library)

1 Lewis Research Center

1 National Science Foundation (Mathematics Division, A. Grad)

1 Center for Naval Analyses, University of Rochester, Arlington, Va.

1 Jet Propulsion Laboratory, CIT, Pasadena, Calif. (Technical Library)

1 Naval Warfare Research Center, Stanford Research Institute, Menlo Park, Calif.

1 Stanford University, Department of Operations Research, Stanford, Calif.  
(Frederick Hillier)